**UNIVERSIDAD COMPLUTENSE DE MADRID**
FACULTAD DE CIENCIAS MATEMÁTICAS


# TESIS DOCTORAL

# Mathematical models for introduction, spread and early detection on infectious diseases in veterinary epidemiology

MEMORIA PARA OPTAR AL GRADO DE DOCTOR

PRESENTADA POR

**Eduardo Fernández Carrión**

Advisors

**Ángel Manuel Ramos**
**Benjamín Ivorra**
**Beatriz Martínez López**

**Madrid, 2018**

**© Eduardo Fernández Carrión, 2017**

UNIVERSIDAD COMPLUTENSE MADRID

# Mathematical models for introduction, spread and early detection of infectious diseases in veterinary epidemiology

Modelos matemáticos de introducción, propagación y detección temprana de enfermedades infecciosas en epidemiología veterinaria


Eduardo Fernández Carrión

*Advisors:*

Dr. Ángel Manuel Ramos
Dr. Benjamin Ivorra
Dra. Beatriz Martínez López


Thesis submitted in fulfillment of the requirements for the degree of
**Doctor of "Mathematical and statistical methods for data processing"**

Tesis presentada en cumplimiento de los requisitos para el grado de
**Doctor en "Métodos estadístico-matemáticos y computacionales para el tratamiento de la información"**

Madrid, April 23, 2017

# Agradecimientos

Las primeras líneas de esta tesis doctoral quiero que sean para expresar mi enorme agradecimiento a mis directores: Benjamín Ivorra, Ángel Manuel Ramos y Beatriz Martínez López. En este orden les conocí. Uno a uno, me introdujeron en un mundo completamente nuevo para mí, la epidemiología. Comencé esta tesis sin saber qué es lo que podía hacer un matemático para ayudar a contener la propagación de enfermedades. Y ahora, al finalizarla, me pregunto por qué no hay más que lo hagan. Han sido muchos años desarrollando ideas, modelos, aplicaciones y escribiendo artículos. Gracias a vuestro apoyo, tanto humano como científico, esta tesis ha podido salir adelante.

Gracias también a José Manuel Sánchez-Vizcaíno. Por la motivación, la ilusión contagiosa y por tener tantas ideas, esta tesis también te corresponde. Gracias por la confianza depositada estos años y por tu incondicional apoyo.

$$MAT + VET = SUCCESS$$

A mis compañeros de matemáticas María, Silvia, Jorge, Luis, Sergio, Marcos y Carlos. Al equipo de Visavet, en especial a la *IT crowd*. Al equipo del CISA, en especial a Ana, Irene, Jaime y María Jesús. A mis compañeros de veterinaria Cristina, Belén, Raquel, Marta, Almudena, Marina, Victor, Elvira, Mar, Chelo, Rocío, Goyache, José Ángel, Sandra, Cecilia y Lina (SUAT). Podría dedicar una tesis entera para explicar los buenos momentos que hemos vivido juntos.

Quiero agradecer en general a todos mis amigos y, en especial, a Oleg por ayudarme con algunas traducciones de esta tesis. Y no me puedo olvidar de Laura que, entre otras cosas importantes, también ha aportado su granito de arena en esta tesis.

Finalmente, a los que me han apoyado toda mi vida...

*...a mi familia.*

# Contents

# Summary

Epidemiology is a discipline that studies the distribution and the factors related to the spread of (infectious) diseases in susceptible populations, and then uses this knowledge to control them. In this discipline, mathematical modeling plays a fundamental role in explaining the influence of potential risk factors, patterns of disease spread and the effects produced.

Applying mathematical models to epidemiology requires identifying the variables that influence the behavior of a disease in order to obtain a set of equations whose solution agrees with what happens in reality. The goal is to generate epidemiological models that not only explain the spread of diseases based on past experimental information, but also predict possible future scenarios. In this way, mathematical modeling is often the basis for decisions and practices aimed at preventing, controlling or eradicating disease.

For the construction of epidemiological models, there is a wide variety of mathematical and statistical techniques. Thanks to the increasingly powerful technologies, there is a great offer of solutions designed for obtaining, processing and analyzing data, even in real time. Together, the epidemiologist can draw on potent platforms to develop new epidemiological tools based on mathematical models that respond to current demand.

The present thesis has focused on the development of three tools that cover surveillance, early detection and decision-making, each aimed at solving a real epidemiological problem in the field of animal health. This thesis is multidisciplinary in its approach and its scope: it combines methodologies in mathematics, statistics and computing to generate models useful to epidemiological and veterinary sciences.

The first tool developed in this thesis is based on a three-level deterministic numerical model to assess the risk of introducing vector-borne diseases by simulating wind-borne transport of insects from infected areas. The first level consists of an advection process, in which insects travel long distances; this process is modeled using a transport-diffusion equation linked to the wind vector. The second level consists of a deposition process, which is modeled using partial derivative equations. The third level consists of insect survival during the journey, which depends on climatic conditions. This three-level model can identify geographical areas and periods of the year associated with higher probability of insect introduction from areas of suspected outbreaks, survival and deposition. This tool is used in this thesis to evaluate

the risk of introduction of bluetongue disease via wind-borne movement of infected Culicoides from North Africa to Spain.

The second tool is a motion-based monitoring system that allows the detection of significant changes in the behavior of animals that develop febrile disease. During the incubation period of such disease, animals gradually reduce their movements, which can be detected using real time video processing that takes advantage of high data processing speeds, high camera resolution and computer vision algorithms. This tool is used in this thesis to detect a reduction in the movement of pigs infected with African swine fever virus; in fact, the automated system detected the reduced movement long before the appearance of the first clinical signs, suggesting that it may improve current disease detection strategies on farms.

The third tool, called Be-FAST (*Between and within Farm Animal Spatial Transmission*), uses Monte Carlo methods to generate different scenarios of disease spread within and between farms based on the integration of stochastic variables. Within farm transmission spread is based on classical compartment epidemiological models. Between farms spread is based on spatial networks of movements. Be-FAST has been integrated into an autonomous computer program to simulate notifiable livestock disease epidemics, which require strict control measures from the authorities. The program can support epidemiological studies to identify areas at greatest risk and the most important routes of diffusion between farms, to evaluate and optimize the efficiency of control measures and to estimate the health and economic impacts of each scenario. Be-FAST is used in the current thesis to analyze two diseases that directly affect the livestock industry: classical swine fever in Segovia (Spain) and Bulgaria, and foot-and-mouth disease in Peru.

The work described in this doctoral thesis has been presented in four publications in first-quartile journals indexed in *Journal Citation Reports* (JCR) as well as two manuscripts that have been submitted to such journals; participation in eight congresses of international prestige; participation in research projects at national, European and international levels; registration of ownership of computer software; and multidisciplinary projects in collaboration with different institutions and international organizations.

# Resumen

La epidemiología es una disciplina que estudia la distribución y los factores relacionados con la propagación de enfermedades infecciosas en poblaciones susceptibles, y que utiliza este conocimiento para controlarlas. En este campo, la modelización matemática juega un papel fundamental para explicar la influencia de los factores que potencian la propagación, los patrones de comportamiento o los efectos que éstos pueden producir.

En pocas palabras, el desarrollo de modelos matemáticos aplicados a la epidemiología consiste en identificar las variables que influencian el comportamiento de una enfermedad para poder obtener un conjunto de ecuaciones cuya solución concuerde con lo que ocurre en la realidad. Así, los modelos epidemiológicos permitirían no solo explicar la difusión de enfermedades basándose en información experimental pasada, sino también predecir posibles escenarios futuros. Por este motivo, la modelización matemática supone en muchos casos la piedra angular para la toma de decisiones destinadas a prevenir, controlar o erradicar cualquier enfermedad.

Para la construcción de estos modelos epidemiológicos, hoy en día se utilizan una gran variedad técnicas estadísticas y matemáticas. Gracias además a las cada vez más potentes tecnologías, existe una gran oferta de soluciones diseñadas para la obtención, tratamiento y análisis de datos, logrando incluso el procesamiento en tiempo real. En conjunto, el epidemiólogo cuenta con potentes plataformas para desarrollar nuevas herramientas epidemiológicas basadas en modelos matemáticos que respondan a la actual demanda.

Así, esta tesis se ha centrado en el desarrollo de tres herramientas que abarcan la vigilancia, la detección temprana y la toma de decisiones orientadas, cada una de ellas, a resolver un problema epidemiológico real en el campo de la sanidad animal. De esta manera, esta tesis no solo muestra un trabajo multidisciplinar que combina las ciencias matemáticas y veterinarias, sino que además combina diferentes metodologías dentro de las matemáticas, la estadística y la informática para el desarrollo de los modelos.

La primera herramienta se basa principalmente en un modelo numérico determinista de tres niveles destinado a evaluar el riesgo de introducción de enfermedades vectoriales mediante la simulación del transporte de insectos arrastrados por el viento desde zonas infectadas. El primer nivel consiste en un proceso de advección, donde los insectos realizan movimientos de larga distancia basándose en una ecuación de

transporte-difusión ligada al vector del viento. El segundo nivel consiste en un proceso de deposición basado en un modelo de ecuaciones en derivadas parciales. Y el tercer nivel consiste en una estimación de supervivencia del insecto durante el trayecto en función de las condiciones climáticas. Los resultados de este trabajo permiten identificar las áreas y períodos del año con mayor probabilidad de introducción, supervivencia y deposición de insectos desde zonas donde previamente existe sospecha de brotes la enfermedad. Concretamente, esta herramienta ha sido utilizada para evaluar el riesgo de introducción de la enfermedad de Lengua azul a través de Culicoides infectados desde el norte de África hasta España.

La segunda herramienta consiste en un sistema de monitorización basado en el movimiento que permite detectar cambios significativos en el comportamiento de los animales cuando están infectados por enfermedades febriles. Las altas velocidades de procesamiento de datos y alta resolución de las cámaras actuales ofrecen algoritmos de análisis de videos en tiempo real. A través de métodos de visión artificial, se ha desarrollado un sistema que permite cuantificar la desaceleración existente en el movimiento de los animales durante el periodo de incubación de este tipo de enfermedades. Para este desarrollo, la reducción en el movimiento de cerdos infectados con el virus de la Peste porcina africana monitorizada mediante estas técnicas fue detectada mucho antes de la aparición de los primeros signos clínicos, mejorando así las estrategias actuales de detección de enfermedades en granjas.

La tercera herramienta, llamada Be-FAST (*Between and within Farm Animal Spatial Transmission*), ha sido integrada en un programa informático autónomo que permite simular casos de epidemias de enfermedades ganaderas de declaración obligatoria (es decir, que exigen estrictas medidas de control por parte de las autoridades). El algoritmo principal está basado en métodos de Montecarlo que generan diferentes escenarios como resultado de la integración de variables estocásticas. Be-FAST utiliza un modelo de difusión de enfermedades dentro de granjas y otro entre granjas. El primero está basado en modelos epidemiológicos compartimentales clásicos de difusión y el segundo en redes de movimientos. Los escenarios resultantes del algoritmo permiten realizar estudios epidemiológicos centrados en identificar las áreas de mayor riesgo, las vías más importantes de difusión entre granjas, evaluar/optimizar la eficiencia de las medidas de control y estimar el impacto sanitario/económico de cada escenario. A lo largo de este trabajo, se han estudiado tres casos de enfermedades que afectan directamente a la industria ganadera: la Peste porcina clásica en Segovia (España) y en Bulgaria, y la Fiebre aftosa en Perú.

El desarrollo de esta tesis doctoral ha permitido la publicación de cuatro artículos científicos y la preparación de dos más presentadas a revistas con un importante factor de impacto, clasificadas en el primer cuartil de su categoría e indexadas en *Journal Citation Reports* (JCR); la participación en ocho congresos de prestigio internacional; la participación en proyectos de investigación a nivel nacional, europeo e internacional; el registro de la propiedad de un software informático; y el trabajo multidisciplinar con diferentes instituciones y organizaciones internacionales.

# Prologue

### Epidemiology in veterinary science

In the last years, investigating disease behaviours in groups of individuals has led the understanding of many problems associated rather than traditional medicine, mainly based on the treatment of individuals with clearly identifiable diseases [146]. Epidemiology is defined as the study of the occurrence and distribution of health-related states or events in specified populations, including the study of the determinants influencing such states, and the application of this knowledge to the control of health problems [118]. The World Health Organization (WHO) defines epidemiology as the study of the distribution and determinants of a disease, and the application of this knowledge to the control of the disease.

Although veterinary medicine is concerned with the epidemiology of animal diseases, human health is inextricably linked to animal health and production, mainly because a number of diseases may be transmitted from animals to humans, and because the link between human and animal populations can be particularly close in regions where animals provide transportation, food or clothing to humans. As a result, any spread of animal disease can have severe economic consequences.

The determinants of animal disease behavior can be physical, ecological, biological, social, cultural or economic. Epidemiology aims to investigate how all these factors may influence the impact of disease in different susceptible populations, in order to guide measures to remove or reduce that impact. To do this, epidemiology seeks to fulfil five sub-objectives [146]: i) determination of the origin of a disease of unknown cause; ii) investigation and control of a disease of unknown or poorly understood cause; iii) acquisition of information on the ecology and natural history of a disease; iv) planning, monitoring and assessment of disease control programs; and v) assessment of the economic effects of a disease, and analysis of the costs and economic benefits of alternative control programs.

Any epidemiological investigation normally involves observation in order to obtain data for posterior qualitative and/or quantitative analysis. Qualitative research draws on non-numerical data to explore disease features mainly for causal hypothesis testing, such as when field observations suggest an association between certain factors and a disease. In contrast, quantitative research involves numerical analysis. Such research covers a wide range of applications and methods, including surveys,

observational-experimental studies, monitoring, surveillance, modeling, disease controlling, and risk assessment. In essence, the various tasks of an epidemiologist could be summarized as i) developing and implementing systems for acquiring, compiling and reporting information; ii) using causal reasoning to understand infectious diseases in global and local contexts; iii) developing statistical analyses to extract epidemiological insights; iv) assessing risk; v) conducting, planning and programming cross-sector specialist research; and vi) communicating.

### The impact of animal diseases

Numerous viral diseases of animals are now known to exist. In general, there are three types of animal diseases according to the occurrence pattern: *endemic diseases*, i.e. those that affect a specific region permanently or periodically through time; *exotic* or *emerging diseases*, i.e. those that usually cross administrative borders into disease-free areas; and *re-emerging diseases*, i.e. those that have been absent or present at low levels in a territory but that cause a significant increase in infections in a short period of time. Diseases can be classified depending on the transmission pathway as *direct-transmission diseases*, i.e. the virus is transmitted from animal to animal via direct contact; or *vector-borne diseases*, i.e. the infection is transmitted via the bite of an infected arthropod species, such as mosquitoes, ticks, bugs, flies or midges. Most animal diseases affect one species in particular, while others may affect more than one, usually triggering different symptoms for each species. Diseases transmitted from animals to humans are called *zoonotic*.

The World Health Organization for Animal Health (OIE) is the intergovernmental organization responsible for improving animal health worldwide and is under the authority of an Assembly of Delegates designated by each member country. In order to maintain animal welfare, the OIE consults with countries to coordinate rapid reporting of new incursions of animal diseases as well as to implement rapid control and eradication measures. Among the large amount of known animal diseases, the OIE focuses on *notifiable diseases*, which are the most relevant to public health in sanitary and economic terms. Rapid response by local authorities is decisive for minimizing the consequences of notifiable diseases. Thus, strict control measures are implemented once a new infection is reported, such as the culling of infected animals, quarantine or trade restrictions.

The potential impact of animal diseases can be complex and generally goes beyond the immediate effects on public health. For instance, any outbreak reported for a notifiable disease may lead to productivity losses for the livestock sector, loss of incomes from activities that depend on animal resources, loss of human well-being, economic losses associated with control measures, devaluation of livestock products on the market, loss of tourism revenue, and political and economic costs of international export/import policies. For instance, severe acute respiratory syndrome triggered a sharp drop in demand in the service sector in 2002-03 throughout southeast Asia; tourism, public transport, retail trade, hospitality and food services were affected as individuals sought to avoid close contact [85]. In 2003 alone, the value

of Canadian beef and cattle exports declined by over $1 billion (US$400 million for beef, $700 million for live cattle) because of control measures and consumer concern over bovine spongiform encephalopathy [64]. The foot-and-mouth disease outbreak in the United Kingdom in 2001 was estimated to have cost the public sector £3 billion and the private sector £5 billion [145]. The World Bank has estimated that an Avian flu pandemic could result in loss of 2% of the world's gross domestic product and cost the global economy $800 billion within one year [156]. In 1997, 429 outbreaks of classical swine fever isolated the Dutch swine industry for 10 months and led to the slaughter of 12 million animals and estimated losses of €2.3 billion [129, 140]. Spain also suffered the consequences of classical swine fever twice, in 1997 and 2001, leading to the slaughter of 1 million animals and the loss of €110 million [45].

These disease occurrences involved extremely fast spread and caused severe sanitary and economic impacts, but they are only the highest-profile examples of the many notifications received by the OIE. In 2016, the OIE received 1,472 notifications and follow-up reports on 63 diseases, 810 of which occurred in Europe. While these outbreaks may not compare with the scale of foot-and-mouth disease or severe acute respiratory syndrome, they usually involve the loss of many animals, which greatly affects local economies.

Most small outbreaks do not grow into large epidemics because of the efficiency of control measures, surveillance systems, epidemiological analysis and coordinated response from the authorities. Nevertheless, despite the best efforts, some outbreaks may spread out of control. This is the case of African swine fever, a disease that has begun to threaten disease-free European countries after starting in Georgia in 2007 and spreading over the Trans-Caucasian countries into the Russian Federation, Ukraine and Belarus [77]. Since then, this disease has been reported in Estonia, Lithuania, Latvia, Moldova and Poland [148, 32]. Avian influenza has shown a similar transboundary spread; in 2016, it affected 22 European Union members. Or bluetongue disease, which infects periodically the south of Europe.

### Epidemiological modeling

As a result of the constant threat to animal health, there is an increasing demand for improving surveillance systems, eradication protocols and disease detection. Epidemiological modeling is a powerful tool for these purposes. Such models are based mainly on direct observations or experimental results subject to many factors that affect its spread: biological factors, such as pathogen mutation or antibody strength; population density; climatic factors, including factors affecting disease carriers; and social, political or economic factors, directly associated with human interests. The basic assumption in epidemiology is that diseases are dynamic and not randomly distributed; as a result, some populations, spatial locations or time periods are more susceptible to disease than others [54]. The epidemiologist must continuously fit the models by exploring the effects of different assumptions, formulations and technologies.

Models, based mainly on mathematical approaches, are simplifications of more complex systems. The goal is to generate epidemiological models that not only explain the spread of diseases based on past experimental information, but also predict possible future scenarios. The art of epidemiological modeling is to make suitable choices in the model formulation so that it is as simple as possible and yet adequate for the question being considered [93]. The epidemiologist can draw from a wide range of mathematical methods to develop models depending on the purpose of the modeling, the amount and quality of data available, and the background and experience of the epidemiologist in generating hypotheses from experimental data that can drive the model. At the same time, the epidemiologist must be aware that the real world is much less precise and predictable than classical mathematical equations, so appropriate tools are needed to deal with uncertainties, imprecision and vagueness [102].

The epidemiologist can draw from numerous modeling types, including compartmental, ecological, climate, and survival models. Models can be based on approaches that are deterministic or stochastic; continuous, discrete in time or non-temporal; non-spatial or spatial; frequentist or Bayesian; homogeneous or heterogeneous with mixed populations; and static or real time. Even more sophisticated epidemiological models are possible [49].

Building mathematical models involves several steps: identification of a problem in the real world, generation of a hypothesis, evaluation of assumptions, data acquisition and analysis, formal development of a model (building systems of equations), model solving, experimentation and validation, and interpretation of results. From the outset, the modeler must have a clear epidemiological target, something that he or she wants to demonstrate based on a previous hypothesis, and the modeler must have access to good-quality data. Otherwise, he or she may waste time on useless or imprecise models. To design a model, the modeler requires experience in mathematics, statistics and computing, as well as knowledge of epidemiology. He or she must consider scalable solutions and maintain clear documentation for future reference, such as when additional or improved data sets become available, or when the model will be integrated with other platforms or technologies. Finally, models must be appropriately validated. This involves ensuring that the model is an adequate representation of the system under study and that its outputs are sufficiently accurate and precise for the intended purpose [93].

### New technologies

Since the 1990s, the rise of the Internet has caused a sudden leap in access to epidemiological data and scientific dissemination. It has probably helped drive technological change that has led to scientific progress in recent years [154]. As a result, new techniques, methodologies and data are easily accessible to a broad audience, accelerating scientific innovation. The Internet facilitates the interdisciplinary sharing of scientific knowledge and applications among mathematicians, statisticians, computer scientists, physicists, and engineers, among others. Databases geograph-

ically remote from one another can share information to process data for scientific and business purposes. The collaborative nature of many epidemiological projects has probably boosted the epidemiology to a large extent.

In recent years, the technological capacity to compute information has grown at 58% per year, and digitized information has increased at 23% per year [59]. This growth indicates excellent potential for data collecting, loading and processing, which means more access to higher-quality data and therefore more accurate epidemiological models. The future of epidemiological monitoring likely lies in the creation of large, real time databases from biosensors, video processors or devices linked by the *Internet of Things* (IoT); such databases will be combined with big data techniques for data mining.

These trends synergize with the proliferation of domain-specific languages for mathematical use and the current generation of programming languages to allow versatile database management, mathematical implementation and friendly-user interface developments. Moreover, new software and libraries have been developed that powerfully enable mathematical methods in epidemiology. The epidemiologist can exploit these platforms to develop new epidemiological tools based on dynamic mathematical models.

### This thesis

The final goal in veterinary epidemiology is rapidly detecting and responding to threats before they become major problems. To accomplish this, the epidemiologist must focus on three essentials: prevent the introduction of diseases into controlled or disease-free areas; detect disease early in the event of introduction; and implement effective control and eradication measures in the event of detection. Epidemiological modeling and other techniques are essential for the epidemiologist to carry out these activities.

Emerging and re-emerging diseases pose a constant threat to the health of animals in disease-free or controlled areas. Disease can be introduced by several pathways: wild animal displacements (e.g. bird migrations), domestic animal movements (e.g. trade in live animals), import or use of animal by-products (e.g. meat), fomites, import of contaminated material, and contamination of animal food. The work of an epidemiologist consists of monitoring all possible introduction pathways and evaluating the risk of this introduction in order to prevent the introduction of target diseases. Thus, epidemiological systems are needed for disease surveillance and for quantifying the risk of introduction. Rapid risk reporting can substantially improve subsequent resource allocation and management for detection and eradication.

Assuming that diseases may be introduced in disease-free areas, the public health authorities should manage systems and protocols for rapid detection. Here, the epidemiologist knows that any infectious disease, in the worst case, may spread indefinitely. In such cases, only the human interaction may avoid full infection in susceptible populations through the deployment of eradication protocols. However,

late detection of infectious diseases delays the deployments of eradication measures and increases the risk of epidemic, followed by an increment in the number of infected animals that may die or may be carriers, i.e. possible sources of new outbreaks. Thus, early detection of infectious diseases is critical. Classical strategies to detect diseases include active surveillance, which involves the repeated collection of data from selected locations to identify changes in animal health or ecological status; passive surveillance, in which local authorities (such as veterinarians) voluntarily report suspicious animal infections; or syndromic surveillance, which involves systematic collection and monitoring of animal mortality in local areas [55]. Technological innovations may support the development of new automatic disease detection systems.

Once a disease has been detected, its eradication depends on the responsiveness of local authorities. In the case of notified diseases, there are action protocols against further spread and eradication. However, poor resource management may lead to costly animal deaths as well as economic losses and, in the worst case, incomplete eradication, allowing the disease to re-emerge or become endemic in the infected area. Effective control and eradication require prior knowledge about the behavior of the target disease, including spread patterns, areas and periods of higher spread risk, economic costs, risk factors and epidemic duration. This information can be obtained directly from the study of specific outbreaks in particular areas, but the challenge of the epidemiologist consists on learning from past events in order to anticipate future disease behavior in the same affected area or in other potential areas with similar characteristics. Epidemiological models or disease simulators can project possible spread patterns onto new locations and times, helping predict possible sanitary consequences as well as the cost-efficiency of possible control measures.

This thesis aims to strengthen collaboration between the veterinary field and mathematics, enlisting the latter for the development of epidemiological tools to accelerate and improve surveillance, decision-making and early detection of diseases in order to mitigate their impact. Thus, this thesis is divided into three chapters describing the development of epidemiological tools for the prevention of infectious disease introduction (Chapter 1), early detection of infectious disease (Chapter 2) and reduction of the risk of epidemic spread (Chapter 3). All three chapters address a current epidemiological challenge involving real cases of animal diseases. All modeling tools were built from different methods in mathematics, statistics and computer science that showed the best fit to the experimental data, making this thesis highly multi-methodological.

In the first chapter, we evaluated disease introduction. Here, we paid special attention to vector-borne diseases due to their emerging importance in European surveillance programmes. Global warming is causing increasing temperatures, which affects the survival and activity of midges and mosquitoes, which transmit infectious diseases. The long-range transport of these vectors by the wind can bring diseases into disease-free or controlled areas, independently of other pathways of introduction, such as trade or infected animal movements. We evaluated the risk of introduction of bluetongue disease into Spain from North Africa via wind-borne transport of biting midges. We collected wind vector field and temperature values from meteorological

stations and we developed a mathematical model based mainly on the advection equation and deposition model with a deterministic solution solved through a Lagrangian approach. The model took into account the effect of temperature on midge survival. We used the model to estimate geographic areas and periods where and when midges were more likely to fall to the ground and thereby come into contact with susceptible livestock. It may be possible to use this model to predict, in real time, the risk of introduction under a probabilistic framework and help create a system of smart active surveillance covering the high-risk areas and periods.

In the second chapter, we evaluated a novel system for disease detection. In this case, we focused on detecting pigs infected with the African swine fever virus. The number of outbreaks of this disease has increased considerably in Eastern European countries since 2007, and the disease is endemic in Sardinia (Italy). Among the clinical signs of this disease, fever causes infected animals to reduce their motion, which we were able to quantify by analyzing video recordings of a herd of pigs experimentally infected with African swine fever virus in an experiment carried out at the Bio-Safety Level 3 (BSL-3) facilities of VISAVET (UCM). Using new video processing technologies and an optical flow algorithm, we monitored animal movements and computed the quantity of animal motion in real time. Furthermore, we used supervised machine learning methods to determine patterns of motion throughout the experiment, and we were able to detect a significant change in animals' daily behavior after infection. It may be possible to extend this technology to the rapid detection of other diseases, which could substantially improve current surveillance protocols, especially on *sentinel farms*, which are farms at high risk of infection that are closely monitored.

In the third chapter, we evaluated the spread of two important diseases in the livestock industry according to three scenarios: i) classical swine fever in Bulgaria, ii) classical swine fever in Segovia (Spain) and iii) foot-and-mouth disease in Peru. We evaluated potential disease spread using the spatial stochastic spread model called Be-FAST, allowing us to test the robustness of the model to different spatial distributions of farms, animal populations, farming uses, resources and policies. Specifically, for the cases i) and ii), the evaluation of an outbreak of the target diseases in these territories had never been explored previously. An appropriate configuration of Be-FAST allowed the precise estimation of spatial risk. In case ii), we implemented a new economic module in Be-FAST in order to compute the economic costs associated with the deployment of risk-based strategies to prevent and control disease. The identification of areas at higher risk of spread and the determination of optimal cost-effective measures to eradicate disease should allow the design of powerful strategies. Be-FAST offers a stochastic implementation of well-known variables related to the farming environment, control measures and outbreaks; the algorithm uses Monte Carlo experiments to predict different output scenarios for later risk analyses and economic estimation. Be-FAST also implements a classical compartmental epidemiological model to estimate within farms transmission, contrary to other disease simulation tools, and for farm-to-farm animal, human and fomites using a spatial network.

In this thesis, each chapter focuses on a different disease, reflecting data avail-

ability and current national and international epidemiological challenges. If the thesis had focused entirely on one disease, it might have been possible to develop a multi-scale system addressing the entire process from infection to introduction and eradication. Nevertheless, the breadth of models, variables and disease behaviors covered in this thesis makes it a valuable foundation-setting reference for approaching various notifiable diseases of significant public health and economic impact. Below are some of the many pending challenges in disease introduction surveillance, early detection and epidemic modeling to which the work in this thesis may contribute.

### Surveillance

Disease surveillance consists of systematic data monitoring in order to detect significant changes in risk factors of disease outbreaks. A disease early warning system combines disease surveillance with analytical tools and epidemiological models for risk assessment within a probability framework [33, 110].

Many early warning systems have been developed to predict impending hazards such as climatic disasters [4, 83], vehicle collisions [86], network hacking [29], financial crisis [22], and student failure [95]. The goal of these systems is to generate alerts of possible hazard events before they occur in order to permit rapid intervention to reduce impact. These systems continually assess indicators that show when hazards may occur imminently: climatic warning systems, for example, extract data quickly from climatic stations or satellites; or vehicle warning systems extract data from integrated sensors to analyze the presence of objects outside. In this way, the first challenge to building an early warning system is to acquire data rapidly, preferably in real time. The second challenge is to process the data rapidly using appropriate predictive hazard models, such as climatic forecasting models [14] or collision detection models [82].

In epidemiological surveillance, there is a wide range of the determinants depending on the target disease. But they are mainly related to ecological factors, as land covers or forestry; climatic factors, as temperature or wind; current location of disease outbreaks; live animal displacements, as live animal trade or wild animal migrations; animal distribution, in farms or wild habitats; animal food imports, as meat trade; vectors distribution, as ticks or mosquitoes suitability maps; or people movements. Disease surveillance systems should be able to monitor changes in all these factors or the quality of epidemiological models could be biased or outdated.

Some types of data, such as those for ecological or climatic factors, are usually of remarkably good quality, with hourly updates and geo-referencing with open access. However, it is unusual to find early warning systems in real time designed specifically for animal diseases. We may have access to wind simulators or temperature forecasts in real time to make reports, but not too many approached or developed for disease surveillance. Sometimes, early warning systems do not need to update data in real time, as the case of ecological factors, which can be updated periodically; and with farm locations and live animal movements, which can be transferred periodically

from administrative databases. However, even these *nearly* real time approaches have not yet to be integrated into most disease early warning systems. Curiously, web-based systems have emerged in recent years to allow text-mining in real time of specific websites for analysis of suspicious disease outbreaks, with excellent results [20, 26, 87].

The role of the environment in the spread and the introduction of infectious diseases in susceptible populations is critical in many cases. Concretely, climatic factors directly affect the abundance and distribution of disease vectors, which should be sampled in the field in order to prove their existence in the area of interest as well as test for disease presence. Direct vector sampling can be complemented by entomological models, which can help identify the most suitable vector habitats in the area of interest as a function of temperature, humidity, vegetation, and proximity to water and food (such as livestock blood). Habitat suitability correlates strongly with vector-borne disease outbreaks. Models that can assess optimal times and locations for vector trapping are valuable for surveillance, especially of re-emerging diseases because they may assess habitat suitability for vector life-cycle or activity; these models rely on classical statistical methods such as logistic regression [36, 23], mixed models [79], random forest [128], and discriminant analysis [73]. Conducting surveillance requires knowledge of introduction pathways, which may be well established in the case of re-emerging diseases, but an automatic monitoring system could be profitable for detecting new incursions.

In the case of small insects such as mosquitoes or Culicoides, numerous studies point to wind currents as the cause of long-distance transportation and therefore the main risk factor contributing to vector-borne disease introduction [56, 157, 126]. However, few published models assess the risk of vector-borne disease introduction through wind currents. Epidemiological models that take wind into account are designed to estimate and predict wind-blow trajectories of gas and dust particles based mainly on advection-diffusion equations. Since dust particles may behave similarly to insects, it should be straightforward to improve these models by adding specific insect characteristics, such as weight or survival capacity at extreme temperatures, yet these models do not include these parameters. One step in this direction is the *Numerical Atmospheric-dispersion Modelling Environment* (NAME) of the UK Met Office, a nearly real time dispersion model to predict wind-blown trajectories of radioactive clouds, chemical accidents, and smoke from volcanic eruptions or fires. The Met Office has partnered with the UK Institute for Animal Health to use NAME to advise on the airborne spread of Bluetongue, foot-and-mouth and Schmallenberg diseases, turning it into an early warning system.

Apart from NAME, most studies of wind-borne trajectories of infected insects are far from being an early warning system. They involve models that are not open-access and that rely on static meteorological data, implying difficult scalability to real time. This has left a substantial gap in active surveillance in many countries. Therefore, these models have been applied mainly to risk analysis of past hazards and to the explanation of past outbreaks; they have been used to predict new outbreaks, but their margin of error is high when climatic variability is also high. The most widely cited model is probably the *Hybrid Single-Particle Lagrangian Integrated Tra-*

*jectory model* (HYSPLIT) [142], developed by NOAA's Air Resources Laboratory. This model simulates atmospheric trajectory and dispersion of wind-blown dust. It has been used to explain the behavior of recent infections or to evaluate the risk of vector-borne disease introduction in different areas [37, 48, 44].

This type of models estimate the spread of gases or particles quite well. However, their main goal is not disease warning. They do not integrate the physical features of the vectors, obliging the epidemiologist to perform additional work for risk assessing, delaying the response time. Furthermore, many studies carried out with public access tools support past and static meteorological input data, hampering real time data integration (and scalability in software). In this work, we propose a model which integrates the physical features of vectors, and that allows an easy numerical implementation for any system able to manage real time climatic data.

### Early detection

Traditional control techniques in classical veterinary medicine focus on the treatment of animals with clearly identifiable disease symptoms. However, the lag between detection of the first physical evidence of disease and official confirmation of infection (generally through blood analysis) delays the application of control measures, which allows disease spread that increases the risk of an epidemic and can cause significant economic losses in affected areas. Classical methods for animal disease detection are based on active, passive and syndromic surveillance. Each of these types of surveillance has disadvantages, since active surveillance can be expensive and labor-intensive, passive surveillance depends directly on the ability of local professionals (e.g. farmers or hunters) to detect clinical signs, and syndromic surveillance may register infections usually after an infected animal has already died.

Most of the notifiable diseases share similar clinical signs, some clearly visible: fever, behavioral changes, skin marks, diarrhea, bleeding and death. Proper training and awareness have helped shorten the lag between first appearance of clinical signs and their detection by veterinarians, farmers and hunters. However, current technologies may reach automatically detection of disease-related changes in behaviors affecting feeding, drinking, body temperature, social behavior, postures and locomotion [103].

Automatization of data monitoring in order to achieve real time surveillance is the future of commercial livestock farming [12]. Simple sensors for water flow or feed supply can report significant reductions in consumption, which may be linked to disease [96]. Infrared video cameras can monitor animal surface temperature, such that continuous monitoring of increasing temperatures can trigger an automatic alert [106, 31]. More sophisticated automated systems acquire digitized readings and relay it for further processing to machine learning or neural network algorithms. For instance, audio recordings of animal coughs, sneezes, screams or groans can be subjected to spectral analysis to identify infected animals [28, 40]. Digital video cameras can convert video data into a computer-readable format for further processing. In fact, the digital processing of images and videos (i.e. sequence of images) has brought

an explosion in applications and algorithms in many disciplines such as medicine, robotics, photography and engineering. As shown by the increasing number of scientific publications based on image processing techniques. This has also increased the number of algorithms designed for specific utilities, and the trend is still upwards. Thus, the emerging of image processing has conducted to its own scientific field of knowledge: the computer vision.

The goal of computer vision is to achieve that a machine can think how a human being would do it. In this sense, the machine receives 2-dimensional images from a 3-dimensional real world. Its challenge is processing image information to figure out objects in different sizes or captured from different points of view, depth, lights, shadows and colors. Working with several algorithms for object recognition and object classification, the machine can identify stationary and moving objects in pictures, including faces (face recognition), tumors (medicine) and vehicles (such as in autonomous vehicles).

In general, object recognition algorithms rely on distinguishing abrupt changes in hue, texture, color or contrast between adjacent pixels on images, which are nothing but a set of sorted pixels. Detection of such abrupt changes allows the algorithm to divide an image into several sets of adjacent pixels; this so-called image segmentation can be performed using one of several techniques, such as k-means or Otsu clustering or Sobel or Canny edge detection. Once segments are identified, they need to be classified as specific objects, e.g. as a human face, a vehicle, a tumor or an animal. Generally, object recognition and classification algorithms rely on machine learning, such as support vector machine or naive Bayes . A database is created containing a reasonable number of images of the same labeled object in different positions, sizes, and colors and with different light effects and shadowing, and the classification algorithm *learns* from this database, allowing it to assess the probability that a new image contains objects in that database, considering possible changes in environment (size or light) and the position of the observer (point of view), which are supported with artificial intelligence algorithms.

In case of video sequences, computer vision focus on tracking of recognized objects, which is very profitable for motion estimation algorithms and predictive behaviour in such objects. Object tracking consist on comparing two consecutive images (or frames) from a video sequence. By applying object recognition algorithms, the same object can be identified in both frames, if an object (or pixel) has change position from one frame to the other, it is possible to estimate the movement, and so the velocity.

Veterinary science and epidemiology have only begun to exploit the tremendous potential of automated object recognition and tracking. Despite of the large potential it shows according to the abilities aforementioned, there are not too many scientific publications in veterinary epidemiology. Some interesting works with veterinary approach are focused on detecting animals through drone-camera records, specifically for wildlife [42]; or studying the animal behaviors in order to identify daily activities as feeding, gather or aggressive [151]; or computing large distant movements [90]. However, the use of this technology for the detection of diseases is

a potential new field of study not too much developed.

Here, we take an enormous step in active surveillance towards the detection of infectious diseases in animals through computer vision. This technology might improve the current surveillance protocols substantially. Thus, it may support the veterinary services in order to take blood test and confirm infection faster and efficiently.

### Epidemic modeling

Another potential entry/spread route for many infectious diseases is the introduction of live animals, by both migratory movements or trade of live animals, or the introduction of animal by-products. Consider for instance, the case of the Newcastle disease in Italy (2000) where a massive importation of female chicks introduced the virus [24]. Or the case of the Porcine reproductive and respiratory syndrome virus outbreak in Switzerland (2012) by the importation of semen [6]. After the exposition to a virus, the incubation period may delay the appearance of clinical signs, making disease detection difficult [62]. However, disease spread is ongoing. These cases may produce, in many situations, epidemics in areas where no previous outbreaks have been reported (or at least recently). This leads to uncertainty in the prediction of the spread evolution, and therefore hinders the quantification of the effects it produces, such as the number of infected farms, the number of culled animals, the duration of the epidemic or the approximate economic costs involved.

There is a need for knowing the potential spread of animal diseases in these regions before it occurs in order to be prepared for response. A variety of tools and methods can be used to assess the epidemic risk of spread. Mathematical models can be particularly useful in situations where there is a wide array of variables relevant to transmission.

On the one hand, classical compartmental models in epidemiology, as SIR, SEIR or SIS, explains disease dynamics of host populations passing through different disease states, such as susceptible, infected or exposed in time (see Brauer et al., [18] for further information). These models are useful for random-mixed populations. On the other hand, taking into account that the spread of disease caused by the movement of animals or animal by-products, social network analysis may allow to compute the epidemic dynamics at the population scale from the individual-level behavior of infections [74]. Each disease has its own characteristics, and, therefore most of them will need a well-adapted simulation model in order to tackle real-life situations [18]. In livestock disease modeling, best fitting models need to be chosen according to farming type in the area of interest. For instance, we may find areas with extensive farming where different groups of animals may be random-mixed, leading classical compartmental method as best fitting model. On the contrary, areas with no contacts between animal farms except live animals trading lead to social-spatial network analysis as best fitting model.

In the literature, we find several models designed for simulation of the potential

spread of livestock diseases in many countries [130, 72, 139, 108, 70, 143]. These models are generally based on spatial-stochastic methods that generate a variety of possible outcomes, allowing the identification of specific geo-referenced regions at risk. However, most of these models focus on between-farm transmission of disease and make unreliable assumptions about within-farm transmission; in addition, they fail to take into account economic assessments. Here, we develop a model that is able to predict the spread of notifiable livestock diseases. It is useful for the estimation of the sanitary and economic consequences under every farming distribution and locations, which may be very profitable for assessing cost-effective control and eradication protocols.

Decision-making in epidemiology requires a careful balance of numerous, sometimes conflicting, factors and interests. Mathematical models can provide data about alternative decisions and their outcomes, but they cannot make decisions on their own. Veterinary epidemiology has begun to draw on decision-making software packages to identify and assess courses of action and provide insights useful for the decision maker and other stakeholders. Such software can help to simultaneously optimize risk management, budget allocation, preservation of social values, compliance with legal requirements and costs of control. It is not about future decisions, but about the future of decisions that must be taken now; that is, about speculative or uncertain aspects of the outcomes of decisions [27]. Epidemiological models of the future will likely need to incorporate an even richer array of decision-making tools and guidance to help professionals navigate the complexities of disease control and eradication.

# OIE notifiable diseases

The World Organization for Animal Health (OIE, `www.oie.int`) is the intergovernmental organization responsible for improving animal health worldwide. Its main objectives are: i) to ensure transparency in the global animal disease situation, ii) to collect, analyse and disseminate veterinary scientific information, iii) to encourage international solidarity in the control of animal diseases, iv) safeguard world trade by publishing health standards for international trade in animals and animal products, v) to improve the legal framework and resources of national Veterinary Services, vi) to provide a better guarantee of food of animal origin and vii) to promote animal welfare through a science-based approach. In 2016, it has a total of 180 Member Countries. Among the formal obligations of these countries is sending information on relevant animal diseases in their territory. To this end, the OIE establishes a single *list of notifiable diseases* for terrestrial and aquatic animals specifically selected due to the high sanitary and economic consequences that they produce. In 2016, it has a total of 118 diseases.

Since this thesis has based its applications on a selection of the most important diseases listed by the OIE, we have developed a very brief description of target diseases (based on the OIE divulging information) for easily reading in the following chapters.

Classical swine fever (CSF) is a contagious viral disease of domestic and wild swine. CSF is found in Central and South America, Europe, Asia and parts of Africa. The most common method of transmission is through direct contact between healthy and infected animals. The virus is shed in saliva, nasal secretions, urine, and feces. Contact with contaminated vehicles, pens, feed or clothing may spread the disease. Pigs can become infected by eating CSF-infected pork meat or products. It has been proven that in parts of Europe, the wild boar population may play a role in the epidemiology of the disease. The disease has been spread through legal (and illegal) transport of animals and by infected swill feeding. In the acute form of the disease, there is fever, huddling of sick animals, loss of appetite, dullness, weakness, conjunctivitis, diarrhea and an unsteady gait. Several days after the onset of clinical signs, the ears, abdomen and inner thighs may show a purple discoloration. Animals with acute disease die within 1-2 weeks. Severe cases of the disease appear very similar to ASF. Because the clinical signs are not exclusive to CSF and vary widely. Laboratory tests are required to detect antibodies or the virus itself [30].

African swine fever (ASF) is a highly contagious hemorrhagic disease of domestic and wild swine. ASF is generally prevalent and endemic in countries of sub-Saharan Africa. In Europe ASF remains endemic only in Sardinia (Italy) but, since 2007, continuous outbreaks of ASF have been reported in Georgia, Poland, Russia or Belorussia. Pigs usually become infected by direct contact with infected pigs or by ingestion of garbage containing unprocessed infected pig meat or pig meat products. Biting flies and ticks (*Ornithodoros moubata*), contaminated premises, vehicles, equipment or clothing can also spread the virus to susceptible animals. Severe cases of the disease are characterized by high fever and death in 2-10 days on average and the mortality rate can be as high as 100%. Other clinical signs may include loss of appetite, depression, redness of the skin of the ears, abdomen, and legs, respiratory distress, vomiting, bleeding from the nose or rectum and sometimes diarrhea. Abortion may be the first event seen in an outbreak. Moderately virulent forms of the virus produce less intense symptoms though mortality can still range from 30-70%. ASF may be suspected based on clinical signs and confirmation must be made through prescribed laboratory tests, particularly to differentiate this disease from CSF [1].

Foot-and-mouth disease (FMD) is a severe, highly contagious viral disease that affects cattle, swine, sheep, goats, and other cloven-hoofed ruminants. All species of deer and antelope as well as elephant and giraffe are also susceptible. Currently, FMD is endemic in several parts of Asia, most of Africa and the Middle East. In Latin America, the majority of countries applied zoning and are recognized free of FMD with or without vaccination, and the disease remains endemic in only a few countries. The significance of FMD is related to the ease of virus spread through new animals carrying the virus (saliva, milk, semen, etc.) may introduce the disease to a herd. Also contaminated pens, buildings or vehicles used to house and move susceptible animals, contaminated materials, people wearing contaminated clothes, infected meat or animal products and aerosol may spread the virus. The clinical signs are more severe in cattle and intensively reared pigs than in sheep and goats. The typical clinical sign is the occurrence of blisters (or vesicles) on the nose, tongue, lips, oral cavity, between the toes, above the hooves, teats and pressure points on the skin. Other symptoms often seen are fever, depression, hypersalivation, loss of appetite and weight, drop in milk production. If infected with the FMD virus, death can occur in young animals before development of blisters due to damage to the heart muscle caused by the virus. Animals that have recovered from infection may sometimes carry the virus and initiate new outbreaks of disease. The disease may be suspected based on clinical signs with confirmation made through prescribed laboratory tests [46].

Bluetongue (BT) disease is a non-contagious, viral disease affecting domestic and wild ruminants (primarily sheep and including cattle, goats, buffalo, antelope, deer, elk and camels) that is transmitted by insects, particularly biting midges of the Culicoides species. BT has a significant global distribution in regions where the Culicoides is present, including Africa, Asia, Australia, Europe, North America and several islands in the tropics and subtropics. The virus is maintained in areas where the climate will allow biting midges to survive over winter. The insect vector is the key to transmission of BT virus between animals. Vectors are infected with BT

virus after ingesting blood from infected animals. Without the vector, the disease cannot spread from animal to animal. Infected cattle play a significant role in maintaining the virus in a region. Cattle may serve as a source of virus for several weeks while displaying little or no clinical signs of disease and are often the preferred host for insect vectors. In infected sheep, clinical signs vary and can include fever, hemorrhages and ulcerations of the oral and nasal tissue, excessive salivation, and nasal discharge and swelling of lips, tongue and jaw, inflammation of the coronary band and lameness, weakness, depression, weight loss, profuse diarrhea, vomiting, pneumonia and 'blue' tongue as a result of cyanosis. Laboratory tests are required to confirm the diagnosis [15].

Rift Valley fever (RVF) is an acute viral disease that can cause severe disease in domestic animals (such as buffalo, camels, cattle, goats and sheep) and humans. RVF is mainly found in countries of sub-Saharan Africa and in Madagascar. Many different species of mosquitoes are vectors for the RVF virus. Thus, RVF is most commonly encountered close to flourishing mosquito populations. Humans are highly susceptible to the RVF virus and may become infected with RVF by being bitten by infected mosquitoes, through contact with blood, other body fluids or tissues during killing, skinning and cutting of infected animals, or by consumption of raw milk or uncooked meat from infected animals. Clinical signs depend on the species of animal affected and conditions such as age and pregnancy. During epidemics the occurrence of numerous abortions and mortalities among young animals, together with disease in humans, is characteristic. Pregnant sheep and cattle affected by this disease will almost always abort (80-100%). People with RVF will either show no symptoms or develop a mild illness. In areas where the disease is known to occur, RVF may be suspected based on clinical signs, insect activity, concurrent disease in animals and humans, rapid spread of the disease and concurrent contributing environmental factors but laboratory tests are required to confirm the diagnosis[124].

# Chapter 1

# Risk of introduction of vector-borne diseases

## Article submitted to peer reviewed journal

- **E. Fernández-Carrión**, B. Ivorra, B. Martínez-López, A. M. Ramos, J.M. Sánchez-Vizcaíno. "*An advection-deposition-survival model to assess the risk of introduction of vector-borne diseases through the wind: application to bluetongue outbreaks in Spain*".

## Proceedings

- M. Martínez-Avilés, **E. Fernández-Carrión**, A. Sánchez-Gómez, C. Amela, J. Lucientes, J.M. Sánchez-Vizcaíno. "*One health framework for the evaluation of Rift valley fever risk of introduction*". 8th Annual Meeting of EPIZONE (23-25 September 2014), Copenhagen (Denmark). Poster.

## 1.1 Introduction

Introduction of vector-borne diseases can substantially harm public health and animal health, causing significant sanitary and financial loss. Long-range wind-borne transportation of infected flying insects has previously been linked to the introduction of viruses affecting humans and/or animals such as West Nile virus, dengue or Rift Valley fever. Therefore modeling the movement of wind-borne midges or mosquitoes may help identify geographic regions and seasons at higher risk of incursion, improving surveillance efforts.

Bluetongue (BT) is a viral disease of ruminants transmitted by biting midges of the genus Culicoides. This virus has traditionally been considered endemic to tropical and subtropical regions, between parallels 35°S and 40°N; however, recently several outbreaks have occurred further North and it is considered to be a re-emerging disease in new latitudes [158]. So far, the disease has been reported in Australia, USA, Africa, Middle East, Asia and Europe although the geographic distribution of the 27 virus serotypes differs by region [94].

Previous studies suggest that Culicoides could be transported by the wind to distances of up to 170 km from their original location [53]. Wind-borne movement of BT virus-infected Culicoides is thought to explain the northward spread of BT in Europe, primarily through routes from Turkey to Greece-Bulgaria, from Algeria-Tunisia to Italy and from Morocco-Algeria to Spain [157]. Specifically, wind-borne introduction of Culicoides to Spain from North Africa has been suggested as the cause of the numerous BT outbreaks in Spain in 2004-05 [34, 100, 5]. In October 2004, 20 BT outbreaks were reported in Spain; this number increased to 298 in December 2004 and to 328 at the start of 2005 [67, 8, 9].

The goal of this work was to develop a model that could identify geographic areas and time periods at highest risk of introduction of small insects potentially infected with human and/or animal pathogens. The model takes into account three processes: (1) advection, in which insects are considered as particles in the air, and they move according to wind currents; (2) deposition, in which wind-borne insects are deposited on the ground; and (3) survival, reflecting insects' survival or mortality as a function of climatic factors. This model was numerically implemented to simulate wind-borne Culicoides introduction from North Africa into Spain and validated using field data on BT outbreaks reported in Spain in 2004.

## 1.2 Materials and methods

### 1.2.1 Culicoides characteristics

Culicoides are 1-3 mm in size, they weigh less than 1 mg and, under favorable conditions, they can be found as high as 200 m above the ground [133]. Previous studies consider that the long-range transportation of Culicoides is likely caused by

wind currents at high altitudes [157, 58, 107]. Therefore the model assumes that the movement of small insects in the sky shows a similar behaviour to wind-borne dust particles [89]. Then, the model takes into consideration the following processes:

(1) Wind advection. Particle movement in the sky due to wind currents is computed using an advection partial differential equation (PDE), adapted here to the specific characteristics of Culicoides. The same kind of PDE has been used to predict the movements of dust [65, 144], sediments [57, 155], oil [2, 69] or gaseous substances [160, 21].

(2) Vertical deposition. The model assumes that Culicoides deposit onto the ground from the sky in a manner similar to Saharan dust particles in the air. Thus, Culicoides deposition is modeled using the Eta/NCEP model of Nickovic et al. in [112] (BSC-DREAM8b, `www.bsc.es`), originally developed to estimate the concentration of Sub-Saharan dust deposited on the ground during transport from Africa to Europe.

(3) Survival. Dead insects do not pose an introduction risk. Consequently, the model estimates a mortality rate during wind-borne transport on the basis of climatic variables.

### 1.2.2 Mathematical model

We consider the spatial domain $\Omega \subset \mathbb{R}^3$ and time interval $[0, T]$. We denote by $\Gamma_{\mathrm{g}}$ and $\Gamma_{\mathrm{s}}(t)$ the subsets of the boundary of $\Omega$ corresponding, respectively, to the ground and to the source of Culicoides entering into $\Omega$ at time $t \in [0, T]$. $C(x, y, z, t)$ is the spatial and temporal distribution of the number density of insects (i.e. the number of insects per unit volume) at point $(x, y, z) \in \Omega$ at time $t \in [0, T]$. Since the number density of insects varies according to wind advection, vertical deposition and survival, the evolution of $C$ is governed by the following equations

$$\begin{cases} \dfrac{\partial C}{\partial t} + \operatorname{div}(C\mathbf{w}) = -\sigma(\mathbf{f}, C) & \text{in } \Omega \times (0, T], \\ C = C_0 & \text{in } \Omega \times \{0\}, \\ C = C_s & \text{on } \{(x, y, z, t) : (x, y, z) \in \Gamma_{\mathrm{s}}(t)\}, \end{cases} \tag{1.1}$$

where $C_0$ is the initial distribution of the number density of Culicoides in $\Omega$, $C_s(x, y, z, t)$ is the number density of insects at point $(x, y, z) \in \Gamma_{\mathrm{s}}(t)$ at time $t \in [0, T]$, $\mathbf{w} = (w_x, w_y, w_z)$ is the velocity field related to the wind and deposition effects, which satisfies $w_z = 0$ on $\Gamma_{\mathrm{g}} \times (0, T]$, and $\sigma(\mathbf{f}, C)$ is the mortality function (i.e. the number of Culicoides that die per unit volume and unit time), which depends on $C$ and the spatio-temporal distribution of climatic factors $\mathbf{f}$ associated with survival.

### 1.2.3 Model implementation for the case of Culicoides

Here, the model formulated above is tailored for studying the case of the introduction of Culicoides from North Africa to Iberian Peninsula, then the model is

validated in numerical experiments.

### Data source

In order to model the wind velocity vector $\mathbf{w}$ and the climatic factors $\mathbf{f}$ in the domain $\Omega \times [0, T]$, wind and temperature data for Spain were provided by the Spanish Meteorological Agency (`www.aemet.es`), while data for Morocco and Algeria were obtained from Weather Underground website (`www.wunderground.com`). Data were daily measurements of temperature, horizontal wind direction (i.e. $w_x$ and $w_y$) and wind speed taken in 2004 at meteorological stations throughout the study area.

Vertical velocity $w_z$ was assumed to be governed by the same processes affecting the deposition velocity of Saharan dust particles [112]. Therefore we used the vertical component $w_z = |(w_x, w_y, 0)| - v_g$ for wind vector $\mathbf{w}$, where $v_g = \frac{2g\rho R^2}{9\nu}$ is the gravitational settling velocity calculated from the Stokes formula after assuming that the midge density is $\rho = 0.75 \frac{M}{\pi \cdot R^3}$; midge radius, $R = 0.5 \cdot 10^{-3}$ m; midge weight, $M = 0.5 \cdot 10^{-6}$ kg; air viscosity, $\nu = 1.78 \cdot 10^{-5}$ $\frac{\text{kg}}{\text{m} \cdot \text{s}}$; gravity acceleration at altitude $z$, $g(z) = g_g(1 - 2\frac{z}{R_T})$; gravity acceleration on the ground, $g_g = 9.81$ $\frac{\text{m}}{\text{s}^2}$; and Earth's radius as $R_T = 6,371$ km [88].

According to the literature, while long-range transportation of Culicoides is caused by wind currents at high altitudes, favorable temperatures and wind speeds for natural Culicoides outdoor mobility are often located below 200 m above the ground [133, 157, 58, 107]. Thus, the model assumes that below this altitude, the Culicoides do not move through wind currents (i.e. $\mathbf{w} = (0, 0, 0)$ if $z < 200$ m).

Temperature values at higher altitudes were estimated using the standard free atmospheric lapse rate [84], $\gamma = -\frac{d\theta}{dz}$, of 6.5°C decrease per kilometer of increasing elevation estimated for Mediterranean atmosphere.

Data on the number of Culicoides trapped per day along the Spanish territory between 2004 and 2013 and average temperatures associated with each capture were collected from the Ministry of Agriculture, Fisheries, Food and Environment (`www.mapama.gob.es`).

### Mortality function

The mortality function predicts the number density of insects based on temperature, since this is one of the most important determinants of Culicoides life cycle and lifespan [52, 135]. Although other climatic parameters, such as relative humidity and precipitation, can also affect insect survival, neglecting these factors only moderately affects the accuracy of model predictions [19]. In order to simplify the model, we only considered the impact of temperature in the Culicoides mortality function $\sigma$; i.e. $\mathbf{f} = \theta$, with $\theta(x, y, z, t)$ being the temperature value (here in °C) at point $(x, y, z) \in \Omega$ and at time $t \in [0, T]$. Therefore the mortality function is defined

by

$$\sigma(\theta, C) = \begin{cases} \mu(-6)C & \text{if } \theta < -6, \\ \mu(\theta)C & \text{if } -6 \leq \theta \leq 42, \\ \mu(42)C & \text{if } \theta > 42, \end{cases} \tag{1.2}$$

where $\mu$ is the Culicoides mortality rate (i.e. the inverse of lifespan) as a function of temperature. This function was derived based on studies reporting total Culicoides mortality after 2 h at 42°C or -6°C [150] and Culicoides lifespan was assumed to be shorter than 24 h after exposure to 40.5°C or -4°C [150]. Culicoides lifespan of 27.5, 18.8, 13.4 and 10.2 days were observed at temperature of 10°C, 15°C, 20°C and 30°C, respectively [159]. These values were used to estimate intermediate values for arbitrary temperatures through the Hermite cubic interpolation computed with the Matlab function *pchip* and denoted by $1/\mu(.)$. Thus the inverse function $\mu(.)$ returns the estimated mortality rate of the Culicoides depending on the value of the temperature (Table 1.1).

| Temperature (°C) | -6 | -4 | 15 | 20 | 25 | 30 | 40.5 | 42 |
|---|---|---|---|---|---|---|---|---|
| Lifespan (days) | 1/12 | 1 | 27.5 | 18.8 | 13.4 | 10.2 | 1 | 1/12 |

Table 1.1: Estimated mean lifespan of Culicoides at different temperatures, based on field data [159, 150].

### Activity function

Field data on the mean number of Culicoides trapped in Spain at temperatures between 0°C and 35°C were fitted to a polynomial function computed with the *polyfit* command in Matlab, which was defined as the activity function and denoted by $\delta(.)$. We used this function to estimate the outdoor number density of Culicoides in source locations, i.e. $\Gamma_s(t)$ (see next Section), for arbitrary temperatures (in °C).

### Numerical implementation

We discretized $\Gamma_g$ into a two-dimensional meshgrid of $n_x = n_y = 50$ (i.e. 2,500 cells), where $n_x, n_y \in \mathbb{N}$. This grid covered the area extending from latitude 34°N to 44°N and from longitude 10°W to 4.25°E, comprising parts of Spain, Portugal, France, northern Morocco and northern Algeria (Figure 1.1).

Since the number density of wind-borne particles is generally greater at lower altitudes than at higher ones [137, 149], we assumed the center of mass of wind-borne living midges to lie at an altitude of 1 km. We also discretized the temporal domain $[0, T]$ into $N \in \mathbb{N}$ time steps using a fixed time interval $\Delta t \in \mathbb{R}$ given by $\Delta t = T/N$. Here, $\Delta t = 1$ day.

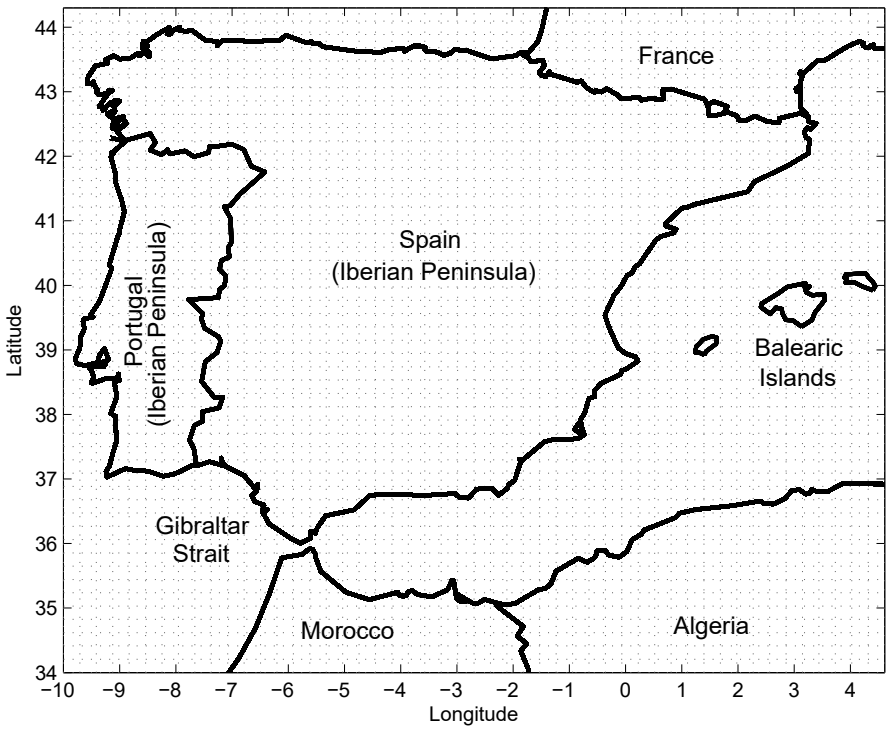


Figure 1.1: Domain $\Gamma_g$ in numerical experiments, extending from latitude 34°N to 44°N and from longitude 10°W to 4.25°E. Black lines demarcate country boundaries.

The system of Equations (1.1) was solved numerically using a Lagrangian particle approach. The source of the Lagrangian particles remained constant (i.e. $\Gamma_s(t) = \Gamma_s$, $\forall t \in [0, T]$) and was defined as lying in $M \in \mathbb{N}$ fixed cells at an altitude of 1 km over northern Morocco and/or northern Algeria. At each time step, $M$ new particles started from $\Gamma_s$, such that the model included $(k+1) \times M$ particles at time step $k \in \{0, ..., N-1\}$, and the total amount of particles per simulation was $N \times M$. All particles were assigned with an index $i \in \{1, ..., N \times M\}$. Therefore, at time step $k$, $M$ particles with indexes $\{k \times M + 1, \cdots, (k+1) \times M\}$ started from $\Gamma_s$.

Particle trajectories were estimated by considering the velocity field $\mathbf{w}$ through the explicit scheme:

$$X^i_{n+1} = X^i_n + \Delta t \mathbf{w}^i_n, \tag{1.3}$$

where $X^i_n$ is the position of the Lagrangian particle with index $i = \{1, ..., (n+1) \times M\}$ at time step $n = n(i), .., N$, $\mathbf{w}^i_n$ is the wind velocity vector at position $X^i_n$ at time step $n$ and $n(i) = \lfloor \frac{i-1}{M} \rfloor$ (i.e. $n(i)$ is the greatest integer less than or equal to $\frac{i-1}{M}$). Then, the number density of Culicoides transported at step $n$ by particle $i$ (which was located at $X^i_n$) was denoted as $C^i_n$.

Given the difficulty of estimating the number of Culicoides placed in a specific location, several epidemiological works assume an initial ratio of mosquitoes or midges instead [11, 125]. In our case, the number density of Culicoides transported by particle $i$ situated in $\Gamma_s$ at step $n$ was assumed to be proportional to the Culicoides activity function, $\delta(.)$, as follows

$$C_n^i = \delta(\theta(X_n^i)), \text{ if } X_n^i \in \Gamma_s. \tag{1.4}$$

Next, the evolution of this particle number density function was governed by the following scheme:

$$C_{n+1}^i = C_n^i - \frac{\Delta t}{2}\left(\mu(\theta(X_n^i))C_n^i + \mu(\theta(X_{n+1}^i)C_{n+1}^i)\right), \tag{1.5}$$

where $i = \{1, ..., (n+1) \times M\}$ and $n \in \{n(i), .., N-1\}$.

According to the features of the Culicoides, the number density of Culicoides deposited on $\Gamma_g$ was calculated at each time step $n$ as the sum of all particles below 200 m altitude; consequently, Equation (1.5) was not applied to such particles. Thus, we denoted by $C_{g,n}^{k,l}$ the ground number density of Culicoides in the cell with index $\{k, l\}$, $k = 1, ..., n_x$ and $l = 1, ..., n_y$, as the sum of the number density of insects of each Lagrangian particle deposited in this cell at step $n$. For mapping purposes, the ground number density of Culicoides between cells $\{k, l\}$ were interpolated through the Spline tool, which minimizes overall surface curvature, and then classified in ten levels by the Jenk's natural breaks tool, where 1 was the lowest value and 10 the highest one, both implemented in ESRI ArcGIS v.10.1 software.

### Simulations

Two numerical experiments were performed to evaluate the risk of Culicoides introduction in Spain from North Africa and thereby allow validation of our model against field data of BT outbreaks reported in 2004. The first experiment compared the introduction of Culicoides along an entire year from the cells located at 1 km above the ground of Morocco alone, Algeria alone, and both countries (i.e. the choice of three different sources $\Gamma_s$). Each simulation started on the first day of 2004 and finished after $T = 365$ days. The goal of this experiment was to identify areas at high risk of midge introduction during one year from different African countries.

The second experiment examined risk of wind-borne Culicoides introduction from both Morocco and Algeria for each month of a year; thus, 12 simulations were performed, each starting on the first day of every month and finishing after T=30 days. The goal of this experiment was to identify months associated with higher risk of Culicoides introduction.

Finally, the results of the numerical experiments were compared with field data from the BT outbreaks in Spain in 2004. This allowed us to validate the model.

## 1.3 Results

### 1.3.1 Mortality and activity functions

Figure 1.2 shows the lifespan associated to different values of temperature (blue dots) displayed in Table 1.1 and the cubic Hermite interpolation computed for intermediate values (red line). Thus, the mean lifespan of Culicoides is

$$\frac{1}{\mu(\theta)} = \sum_{i=0}^{3} a_k(\theta)(\theta - b(\theta))^k, \tag{1.6}$$

where the value of $a_k(\theta)$, for each $k \in \{0, ..., 4\}$, and $b(\theta)$ are given in Table 1.2, depending on the value of $\theta$.

| $\theta$ range (°C) | $a_0$ (days) | $a_1$ (days/°C) | $a_2$ (days/°C$^2$) | $a_3$ (days/°C$^3$) | $b$ (°C) |
|---|---|---|---|---|---|
| -6 $\leq \theta \leq$ -4 | 1/12 | $3.69{\cdot}10^{-1}$ | $1.48{\cdot}10^{-2}$ | $1.49{\cdot}10^{-2}$ | -6 |
| -4 $\leq \theta \leq$ 15 | 1 | $6.07{\cdot}10^{-1}$ | $1.56{\cdot}10^{-1}$ | $-6.04{\cdot}10^{-3}$ | -4 |
| 15 $\leq \theta \leq$ 20 | 27.5 | 0 | $-7.77{\cdot}10^{-1}$ | $8.59{\cdot}10^{-2}$ | 15 |
| 20 $\leq \theta \leq$ 25 | 18.8 | -1.33 | $4.58{\cdot}10^{-2}$ | $9.40{\cdot}10^{-4}$ | 20 |
| 25 $\leq \theta \leq$ 30 | 13.4 | $-8.04{\cdot}10^{-1}$ | $8.55{\cdot}10^{-2}$ | $-1.05{\cdot}10^{-2}$ | 25 |
| 30 $\leq \theta \leq$ 40.5 | 10.2 | $-7.40{\cdot}10^{-1}$ | $-6.91{\cdot}10^{-2}$ | $5.11{\cdot}10^{-3}$ | 30 |
| 40.5 $\leq \theta \leq$ 42 | 1 | $-5.89{\cdot}10^{-1}$ | $7.30{\cdot}10^{-2}$ | $-9.06{\cdot}10^{-3}$ | 40.5 |

Table 1.2: Piecewise cubic Hermite interpolating polynomial parameters for the mean lifespan of Culicoides according to temperature values from -6°C to 42°C, see Equation (1.6).

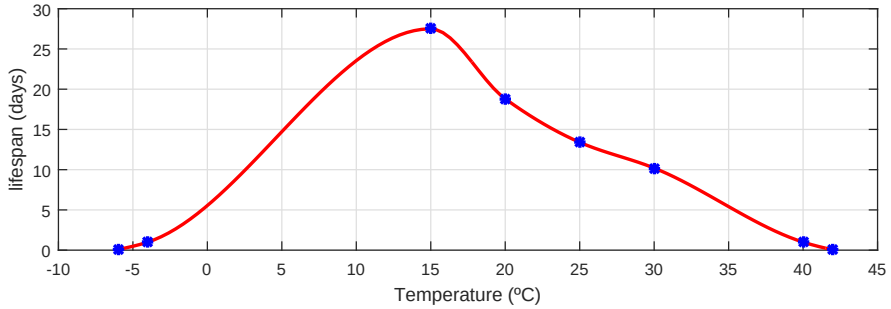


Figure 1.2: Culicoides lifespan used in the present model. The variation of lifespan with temperature is based on field data (blue dots), as well as cubic Hermite interpolation (red line), $1/\mu(.)$.

Figure 1.3 shows the mean number of Culicoides captured (blue dots) and the approximation polynomial fitting (red line). Thus, considering a midges abundance higher or equal to zero, the Culicoides activity function is defined as

$$\delta(\theta) = \sum_{i=0}^{8} c_k \theta^k, \tag{1.7}$$

where $c_0 = -0.226$, $c_1 = 0.344\ ^\circ\text{C}^{-1}$, $c_2 = -0.178\ ^\circ\text{C}^{-2}$, $c_3 = 4.24\ \cdot 10^{-2}\ ^\circ\text{C}^{-3}$, $c_4 = -4.85 \cdot 10^{-3}\ ^\circ\text{C}^{-4}$, $c_5 = 2.89 \cdot 10^{-4}\ ^\circ\text{C}^{-5}$, $c_6 = -9.04 \cdot 10^{-6}\ ^\circ\text{C}^{-6}$, $c_7 = 1.38 \cdot 10^{-7}\ ^\circ\text{C}^{-7}$ and $c_8 = -7.90 \cdot 10^{-10}\ ^\circ\text{C}^{-8}$, if $\theta \in [1.25, 33.45]\ ^\circ\text{C}$; and $c_k = 0\ ^\circ\text{C}^{-k}$, for each $k \in \{1, ..., 8\}$, otherwise.

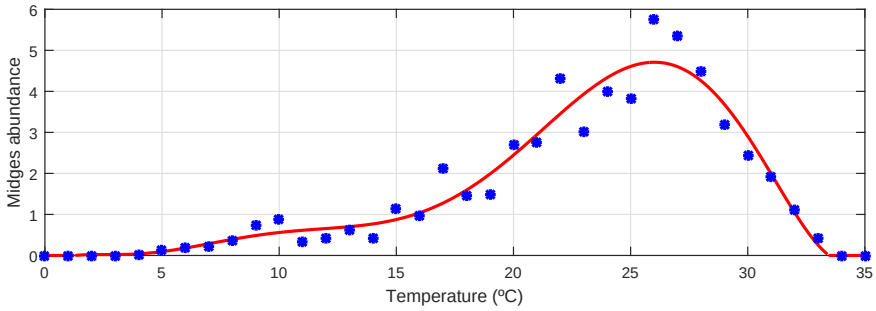


Figure 1.3: Averaged Culicoides trapped per capture (blue dots) and polynomial approximation curve (red line), $\delta(.)$.

### 1.3.2 Annual risk of introduction

Figure 1.4 shows the number density of Culicoides deposited across the entire year on Spanish territory via wind-borne transport from the skies over Morocco, Algeria or both. The most exposed areas were in the south and east of the Iberian Peninsula and Balearic Islands. Midge deposition patterns varied considerably depending on whether the source was Morocco or Algeria. When Morocco was the source, midge depositions reached the 55.4% of the total depositions, and they occurred close to the Strait of Gibraltar. When Algeria was the source, midge depositions reached the 44.6% of the total depositions, and they occurred in a region from southeastern Spain to the eastern coast and Balearic Islands.

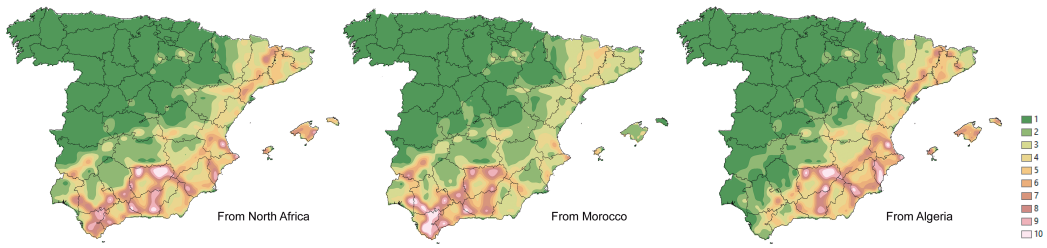


Figure 1.4: Simulated spatial distribution of the number density of Culicoides deposited on the Spanish territory throughout 2004, when the midge source was both Morocco and Algeria (left), Morocco alone (center) or Algeria alone (right). Number densities were binned into levels from 1 to 10 (highest).

### 1.3.3 Monthly risk of introduction

Since changes in wind currents and temperature during the year can affect the dynamics of midge transport from North Africa, we simulated the spatial density of deposited Culicoides over all 12 months of the year. Table 1.3 shows the percentage of the number densities of Culicoides deposited on the ground during each month of 2004. Most deposition occurred between July and September, when number density amounted to 55.5% of the annual total. Approximately equal proportions of depositions occurred between December and May (22.2%) and in the months of June, October and November (22.3%).

| J | F | M | A | M | J | J | A | S | O | N | D |
|---|---|---|---|---|---|---|---|---|---|---|---|
| 3.30 | 4.51 | 5.98 | 1.83 | 2.96 | 8.98 | 14.95 | 23.75 | 16.81 | 6.94 | 6.41 | 3.58 |

Table 1.3: Simulated percentage of number density of Culicoides deposited on the ground during each month of 2004.

Next we examined the spatial distribution of depositions over the 12 months of the year. Figure 1.5 shows the monthly distribution of Culicoides number density deposited on Spanish territory from the source in both Morocco and Algeria. The largest densities of deposition occurred in the southern and eastern parts of the Iberian Peninsula and Balearic Islands, consistent with the annual-level simulation. Risk of wind-borne incursions were high throughout the year in the southern part of the Iberian Peninsula, specifically close to the Strait of Gibraltar; in contrast, this risk was high in the eastern part of the peninsula and on the Balearic Islands only between June and September.

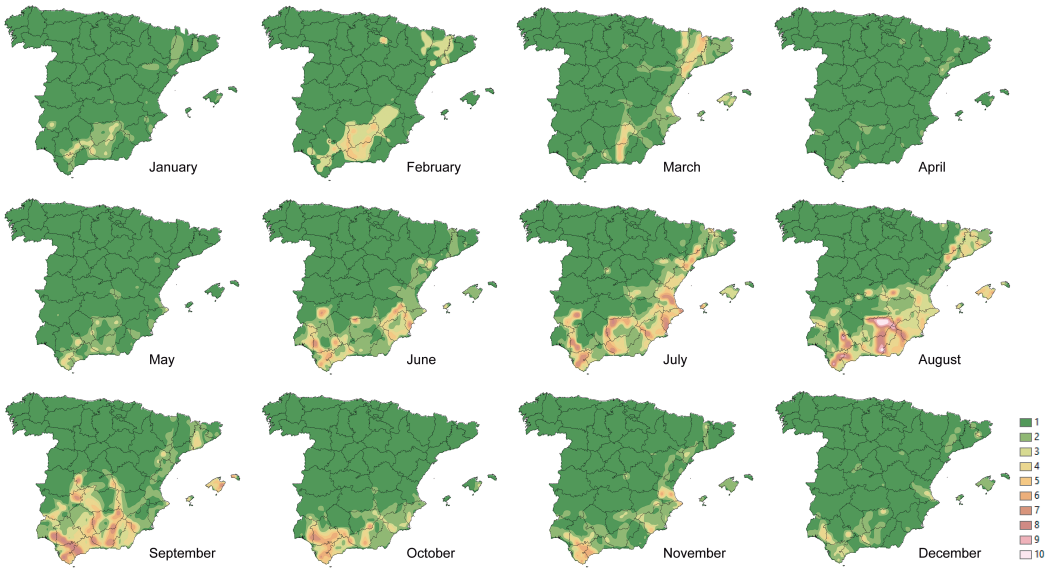


Figure 1.5: Simulated spatial distribution of the number density of Culicoides deposited on Spanish territory for each month of 2004, when the midge source was both Morocco and Algeria. Number densities were binned into levels from 1 to 10 (highest).

### 1.3.4 Model validation

During 2004-05, a total of 328 BT outbreaks were reported in Spain [67, 8, 9]. Most of these outbreaks, 83%, were located in the south of Spain and only 17% were located in the south-west. Figure 1.6 shows the locations of the first 20 BT outbreaks reported in October 2004 on the map of the annual simulation in the same year from North Africa created in Figure 1.4.

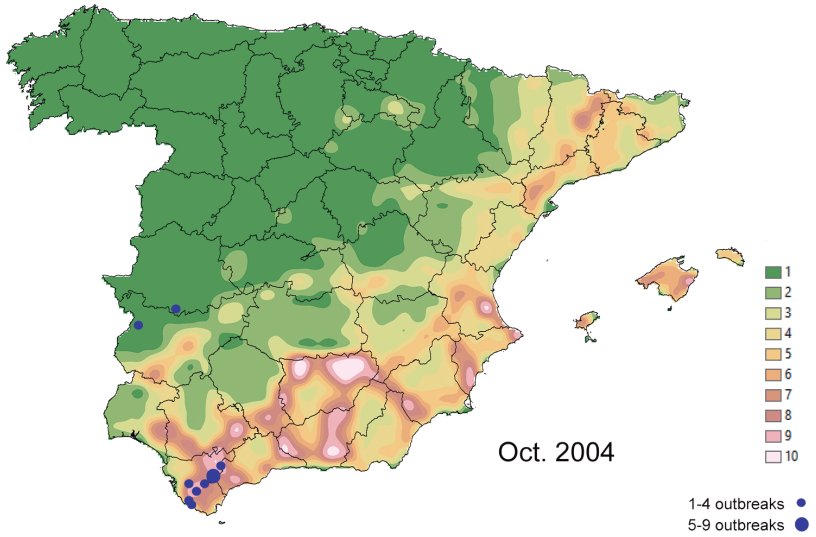


Figure 1.6: Overlap between the first 20 bluetongue outbreaks reported in Spain in October 2004 and the simulated spatial distribution of the number density of Culicoides deposited on Spanish territory throughout 2004, when the midge source was both Morocco and Algeria. Number densities were binned into levels from 1 to 10 (highest).

## 1.4 Conclusions and discussion

In this chapter, we developed a methodology for estimating risk of long-range wind-borne transport and deposition of small insects, while taking into account the likelihood of insect survival given certain climatic conditions. The methodology relies on a mathematical advection model that is solved numerically using Lagrangian particle tracking based on the wind velocity field and the Eta/NCEP deposition equation that takes into account temperature-based insect lifespan and activity. In particular, we focused on assessing the risk of introduction of Culicoides from North Africa to Spain in order to explain the BT outbreaks that occurred in this territory during 2004.

Two numerical experiments were carried out to simulate the incursion of Culicoides through wind currents in 2004 in Spain, one yearly and one monthly. The results showed that the highest number density of Culicoides deposited on regions

agreed with the real primary BT reported outbreaks in the period of study. In particular, most of these locations agreed with suitable areas for Culicoides depositions in the south of Spain. For the remaining reported outbreaks and for future reported ones in 2005 [9], Calvete et al. [23] identified suitable areas for Culicoides based on climatic factors and livestock distribution in Spain. Those regions were distributed mainly in the south and south-west of Spain. The model developed in this work aimed to assess midges' incursions from North Africa due to wind currents, not the spread within the Iberian Peninsula. In future works, this model should be linked with the BT status in North African areas, with the host population distribution and with spatio-temporal suitable conditions for vector survival in order to estimate not only the risk of introduction but also the risk of spread.

Our results suggest that the Culicoides source can substantially affect midge deposition in Spain: when midges came from over northern Morocco, they usually deposited close to the Gibraltar Strait; when they came from over northern Algeria, they usually deposited along southeastern Spain and the Balearic Islands. Our results further showed risk variation during the year: risk of introduction was relatively high between June and September, and moderate during other months. These findings illustrate the power of our approach for helping to focus surveillance efforts in space and time.

Consistent with our modeling, images from the Moderate Resolution Imaging Spectroradiometer Satellite show the large-scale movement of dust from the Sahara Desert and North Africa towards the north and east. Images from 2004 and 2005 show that wind currents in August transported dust particles from over Morocco to southern Spain, and from over Algeria to the southwestern Iberian Peninsula and the Balearic Islands (Figure 1.7). It is therefore entirely plausible that BT virus-infected midges present over Morocco or Algeria at these moments could be carried by wind to Spain.

This work was motivated because, first, different authors foretold the Gibraltar Strait as one of the main BT introduction pathways in Europe [158, 157] and, second, the presence of Culicoides and BT outbreaks were reported along the region and period of study, which allowed the validation of our predictions [67, 8]. The use of the methodology presented here may serve to assess regions and periods at high risk of incursion of several vectors as midges or mosquitoes. Indeed, Culicoides may also be a vector for African horse sickness, and many species of mosquitoes are vectors for diseases such as West Nile virus, dengue, Rift Valley fever, Zika, etc. In future works, real time wind and climatic databases should be embedded in this model in order to launch risk alerts when suspicious initial infected areas may spread the vector to long-range distances, especially where livestock population may be exposed. This development could contribute to an effective early warning system for preventing and controlling vector-borne disease incursions.

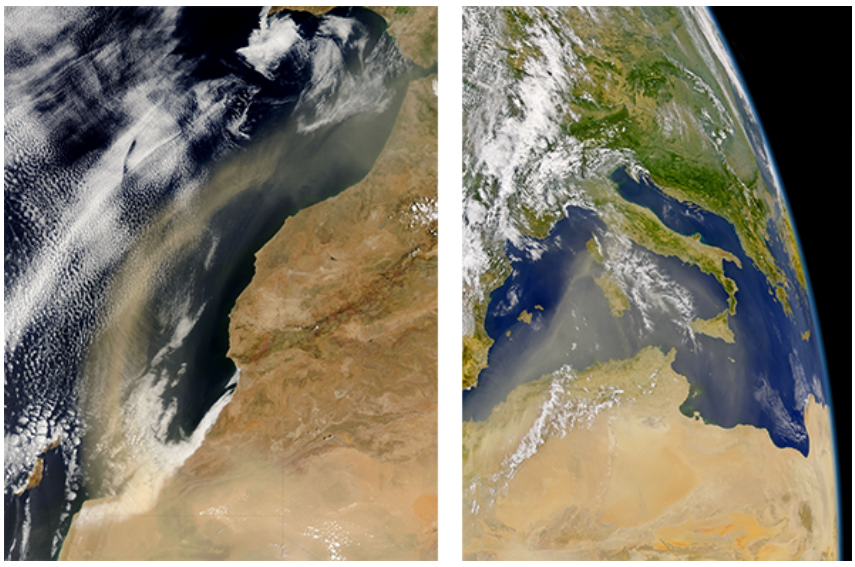

Figure 1.7: Images taken by the Moderate Resolution Imaging Spectroradiometer Satellite over Morrocco on 15 August 2005 (left), showing a dust storm sweeping northward out of Africa; and over the Mediterranean Sea on 19 August 2004 (right), showing a large plume of Saharan Desert moving northeast from Algeria. Source: `www.earthobservatory.nasa.gov/NaturalHazards/view.php?id=15358` and `www.earthobservatory.nasa.gov/NaturalHazards/view.php?id=13715`.

# Chapter 2

# Early detection of livestock diseases

## Article in peer reviewed journal

- M. Martínez-Avilés, **E. Fernández-Carrión**, J.M. López García-Baones, J.M. Sánchez-Vizcaíno. "*Early detection of infection in pigs through an online monitoring system*". Transboundary and emerging diseases (2015).

## Article submitted to peer reviewed journal

- **E. Fernández-Carrión**, B. Ivorra, B. Martínez-López, M. Martínez-Avilés, A.M. Ramos, J.M. Sánchez-Vizcaíno. "*Motion-based video monitoring for early detection of livestock diseases: the case of African swine fever*".

## Proceedings

- **E. Fernández-Carrión**, B. Ivorra, A.M. Ramos. "*Image processing for early detection of livestock diseases*". X UCM Modelling Week (13–17 June 2016), Madrid (Spain). Problem resolution.
- M. Martínez-Avilés, **E. Fernández-Carrión**, I. Bernal-Orozco, B. Rivera, M. Mazariegos, J.M. Sánchez-Vizcaíno. "*Early detection of transboundary animal diseases with a real time monitoring system online*". 8th Annual Meeting of EPIZONE (23-25 September 2014), Copenhagen (Denmark). Poster.

## 2.1 Introduction

Late detection of emergency diseases leads to disease spread, increases the risk of epidemic and causes significant economic losses in affected countries, as demonstrated by the last epidemics in Europe of foot-and-mouth disease [50, 104], classical swine fever [140, 39] or the current epidemic of African swine fever (ASF) [132]. Classical strategies to detect diseases on the farm include active and passive surveillance. A widely used early detection approach is periodic sampling of vulnerable populations (sentinel sampling), which focuses on farms at higher risk of incursion [10].

The effectiveness of passive surveillance relies on the ability to detect disease based on the observation of clinical signs. This can pose a severe obstacle to early detection, particularly if death is the most visible clinical sign as is the case with the aforementioned diseases. On the other hand, farms may apply active surveillance approaches to gain information about disease prevalence on the farm under a sampling protocol based on quantitative polymerase chain reaction (qPCR) of blood samples, which can allow rapid and reliable detection of infection onset. While this can be effective, it is financially and logistically burdensome and not always feasible. This highlights the need for new surveillance methods that can complement sentinel surveillance at a much lower costs, without compromising sensitivity.

In this chapter, we explored a novel system for early detection of ASF based on real time animal monitoring, within the framework of the EU-funded Rapidia Field project (`www.rapidia.eu`). ASF provides an excellent focus for early infection systems for several reasons. It is one of the top-priority animal health diseases in the EU, because since 2007 it has been circulating in Russia and spreading northwards and westwards. ASF has so far reached the EU member states as Lithuania, Poland, Latvia and Estonia. Animals infected with ASF virus rarely recover [105]. Despite the high-fatality rate, transmission is not explosive and clinical signs can range from the non-specific such as fever or prostration to sudden death [131], being fever the fundamental indicator for ASF detection.

Many animal diseases manifest fever and weakness which may be difficult to detect by simple observation in early stages of infection. However, it can be detected through continuous-quantitative monitoring of animal behaviour through new technologies [103]. We aimed to demonstrate that animals infected by ASF show a progressive deceleration in performing daily activities caused by muscle weakness from early stages of infection, which should translate into reduced overall motion, which quantitative video monitoring should be able to detect. This might allow detection of ASF onset in real time. Therefore, we continuously monitored eight pigs housed in a single indoor pen under Biosafety Level 3 (BSL-3) experimental conditions for 11 days before, and then 12 days after, experimental infection with ASF virus. Animal motion was quantified and classified using video processing and supervised learning, respectively, in order to detect significant early changes in behaviour between pre- and post-infection periods.

Latest research and developments based on computer vision techniques have improved the current technology mainly designed for motion detection, object recognition and object tracking. With this work, we have implemented algorithms for motion capture, in which movement recorded by a camera is translated into digital models. It also incorporates a supervised learning method to classify motion-based animal behaviour. Finally, the system aims to assess the randomness of runs in order to detect a reduction in daily motion, such that a certain number of consecutive low motion scores may be attributable to ASF.

## 2.2 Motion-based video monitoring

### 2.2.1 The experiment

The study protocol was approved by the Ethics Committee of the Veterinary Military Centre according to the Spanish Regulations on Animal Welfare (Spanish Royal Decree 53/2013). The experiment was carried out in the BSL-3 facilities of the VISAVET Centre at the Veterinary Faculty of the Universidad Complutense de Madrid as described by Martínez-Avilés *et al.* [101]. The experiment lasted from 11 April 2015 (day 1) to 3 May 2015 (day 23), during which eight healthy pigs (four months old) were enclosed in one indoor pen maintained at constant temperature and humidity, with *ad libitum* access to food and water. They were continuously monitored using one camera in a fixed position. At least twice per day, animals were evaluated by veterinarians and technicians who monitored animal internal temperature, recorded clinical symptoms [117], kept the pen clean and controlled feed supplies. Once per day, blood and oral samples were analysed for ASF virus using a qPCR test [78].

All animals were accommodated in the pen and allowed to acclimate for one day before the experiment. On day 11, all animals were simultaneously inoculated with an attenuated strain of ASF virus from a tick captured in Kenya in 2005 (ASFV Ken05/Tk1). The first positive qPCR detection of the virus occurred on day 15 in six of eight animals. Subsequently, some animals showed mild clinical signs but first evident clinical signs associated to ASF started from day 18 in five of eight animals. After this day, all animals spent increasing amounts of time lying down until ultimately dying from the disease. The last animal died on day 23. During the study, all possible efforts were taken to prevent unnecessary suffering of the animals.

Based on this time course of infection, the experimental period was divided into 4 phases for analysis (Figure 2.1): (1) pre-infection (days 1-11), when animals were free of infection; (2) infection (days 12-15), after animals were inoculated with ASF virus but with no ASF virus detected; (3) qPCR detection (days 16-18), during which ASF virus was qPCR-detected in at least one animal; and (4) clinical detection (days 19-23), when clinical signs of ASF were evident in at least one animal, until the end of the experiment.

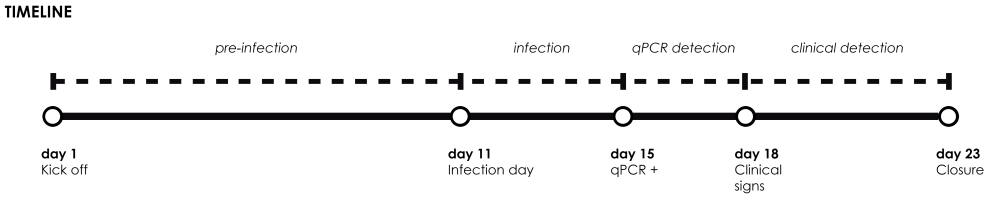


Figure 2.1: Timeline showing the four analytical phases of the experiment. The experiment began on day 1 (kick off), and animals were inoculated with ASF virus on day 11. On day 15, six of eight animals tested positive for the virus by quantitative PCR. On day 18, clinical signs of ASF became evident in five of eight animals. The experiment ended on day 23.

### 2.2.2 Video filming

A total of 552 hours of video footage were taken using a single camera equipped with night-vision capability at a fixed position in an upper corner of the pen (Figure 2.2A). Video footage was collected continuously throughout the 23-day experiment. The room was artificially illuminated from 7 a.m. to 9 p.m., while the camera recorded under night conditions from 9 p.m. to 7 a.m. The 'U'-shape and the limited size of the pen as well as the tall metal enclosure did not allow individual tracking of the animals. Since the fixed camera filmed the same background throughout, and light intensity remained fairly constant within the day and night periods, it was easy to distinguish between animals in motion and static background, despite the standard definition (SD) resolution of the video which might make subsequent motion detection less sensitive.


Figure 2.2: A) Original video frame (all RGB channels). B) The same frame in the red channel. C) Region of interest in the red-channel frame used for motion analysis.

Video was recorded at 6 frames per second (704 x 576 pixels) in RGB24 format, providing 24 bits in red, green and blue channels (RGB) for broad range of colour. We used the red channel for all analyses since it provided the best contrast for animal recognition (Figure 2.2B). In addition, we analysed only a region of interest in which all animals were always observable (Figure 2.2C).

### 2.2.3 Motion capture

Animal movements were recorded and digitally processed through the Optical flow algorithm based on the Horn-Schunck methodology [60] implemented in Matlab. This algorithm estimates the speed and direction of moving objects between consecutive images based on the movement of brightness patterns [51]. This estimation process is described below.

Let us suppose that we know the value of a function $E(x, y, t)$ determining the brightness of any point $(x, y)$ in the (fixed) 2-dimensional domain $D$ recorded in the video, at any time $t \in [0, T_{\max}]$ during recording. Let us consider a point following a trajectory $(x(t), y(t))$ in $D$. If this point maintains its brightness along the entire time interval, the value of $E(x(t), y(t), t)$ remains constant for any $t \in [0, T_{\max}]$. Then, assuming that functions $E$, $x$ and $y$ are smooth,

$$\frac{\mathrm{d}}{\mathrm{d}t} E(x(t), y(t), t) = 0, \forall t \in [0, T_{\max}], \tag{2.1}$$

or, after applying the chain rule for differentiation,

$$\frac{\partial E}{\partial x}(x(t), y(t), t)\frac{\mathrm{d}x}{\mathrm{d}t}(t) + \frac{\partial E}{\partial y}(x(t), y(t), t)\frac{\mathrm{d}y}{\mathrm{d}t}(t) + \frac{\partial E}{\partial t}(x(t), y(t), t) = 0, \forall t \in [0, T_{\max}]. \tag{2.2}$$

In this equation, $(\frac{\mathrm{d}x}{\mathrm{d}t}(t), \frac{\mathrm{d}y}{\mathrm{d}t}(t))$ represents the velocity of the point that at time $t$ is in $(x(t), y(t))$. Since we cover the entire domain $D$ with all possible trajectories, we deduce from Equation (2.2) that

$$\frac{\partial E}{\partial x}(x, y, t)u(x, y, t) + \frac{\partial E}{\partial y}(x, y, t)v(x, y, t) + \frac{\partial E}{\partial t}(x, y, t) = 0, \forall (x, y, t) \in D \times [0, T_{\max}], \tag{2.3}$$

where $(u(x, y, t), v(x, y, t))$ is the velocity of the point located at $(x, y) \in D$ at time $t$. This can be expressed more compactly as

$$E_x u + E_y v + E_t = 0, \quad \text{in } D \times [0, T_{\max}]. \tag{2.4}$$

To compute the velocity function $(u, v)$, the Horn-Schunck method [60] minimises the error $\zeta(u, v)$ given by

$$\zeta(u, v) = \int_D (E_x u + E_y v + E_t)^2 \,\mathrm{d}x\mathrm{d}y + \eta^2 \int_D \left(u_x^2 + u_y^2 + v_x^2 + v_y^2\right) \mathrm{d}x\mathrm{d}y, \tag{2.5}$$

where $\eta$ is a weighting factor that scales global smoothness. Using the theory of calculus of variation, the velocity $(u, v)$ minimising Equation (2.5) satisfies the Euler–Lagrange equations of $\zeta$, given by

$$\begin{aligned} E_x^2 u + E_x E_y v &= \eta^2 \Delta u - E_x E_t, \\ E_x E_y u + E_y^2 v &= \eta^2 \Delta v - E_y E_t, \end{aligned} \tag{2.6}$$

where $\Delta u$ and $\Delta v$ are the Laplacians of $u$ and $v$, defined as

$$\Delta u = \frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} \quad \text{and} \quad \Delta v = \frac{\partial^2 v}{\partial x^2} + \frac{\partial^2 v}{\partial y^2}. \tag{2.7}$$

In reality, the entire video scene consists of $N_{\max} \in \mathbb{N}$ frames (or images) instead of continuous time $t \in [0, T_{\max}]$, $E(x, y, t)$ takes values in $\{0, 1, \cdots, 255\}$ and $D$ is divided into a matrix of 704 x 576 pixels. Therefore, we considered the discrete function $E_{i,j,k} \in \{0, 1, \cdots, 255\}$ as the measured average brightness of the pixel at the intersection of the $i$th row and $j$th column in the $k$th frame. Hence, the minimum $(u_{i,j,k}, v_{i,j,k})$ of Equation (2.5) at an arbitrary pixel $(i, j)$ in the discrete domain $D$ of $704 \times 576$ pixels per frame $k \in \{0, 1, \cdots, N_{\max}\}$ can be estimated iteratively as follows [60]:

$$\begin{aligned} u_{i,j,k}^{n+1} &= \overline{u}_{i,j,k}^{n} - E_x \frac{E_x \overline{u}_{i,j,k}^{n} + E_y \overline{v}_{i,j,k}^{n} + E_t}{\eta^2 + E_x^2 + E_y^2}, \\ v_{i,j,k}^{n+1} &= \overline{v}_{i,j,k}^{n} - E_y \frac{E_x \overline{u}_{i,j,k}^{n} + E_y \overline{v}_{i,j,k}^{n} + E_t}{\eta^2 + E_x^2 + E_y^2}, \end{aligned} \tag{2.8}$$

where $u_{i,j,k}^{0} = v_{i,j,k}^{0} = 0$, $n = \{1, \cdots, N\}$, $N = 25$ is the maximum number of iterations, $\eta = 10$ considering a large relative motion between frames [76], and $(\overline{u}_{i,j,k}^{n}, \overline{v}_{i,j,k}^{n})$ is the neighbourhood average of $(u_{i,j,k}^{n}, v_{i,j,k}^{n})$ computed using the convolution kernel $[1\ \sqrt{2}\ 1; \sqrt{2}\ 0\ \sqrt{2}; 1\ \sqrt{2}\ 1]/4(1+\sqrt{2})$ [71]. We computed $(E_x, E_y)$ using the convolution kernel [–1 –2 –1;0 0 0;1 2 1] and its transposed form for each pixel in the first image and $E_t$ between consecutive images using the [–1 1] kernel [76, 43].

Finally, we computed the global motion in frame $k$ as

$$m_k = \sum_{i=1}^{704} \sum_{j=1}^{576} \sqrt{u_{i,j,k}^2 + v_{i,j,k}^2}. \tag{2.9}$$

Equation (2.9) allows us to obtain a unique motion value for each frame $k$ and to reduce motion analysis to time series analysis.

### 2.2.4 Motion smoothing

We analysed the global motion, $m_k$, in order to detect a significant reduction in animal motion following ASF infection. Nevertheless, two main facts had to be considered before the analysis. First, the video resolution caused a perturbation in the values of $m_k$. Indeed, even when no motion was recorded, we observed that the background slightly changed between consecutive frames causing a problematic baseline noise in the time series. Second, as mentioned previously, the human factor altered the real values of global motion when the workers were in the region of interest of the screen. Thus, some intervals of time scored excessively high values of $m_k$ only when the workers were in the pen.

To smooth the perturbation in $m_k$, we considered a moving average filter by replacing each data point in the time series with the average of the neighbouring data points within the span $K \in \mathbb{N}$. In this work, we considered $K = 45$ in order to average 91 frames; that is, 15 seconds approximately (Figure 2.3). Henceforth, all motion works were carried out throughout the new time series

$$\hat{m}_k = \frac{1}{2K+1} \sum_{i=-K}^{K} m_{k+i}, \forall k \in \{1, \cdots, N_{\max}\}. \tag{2.10}$$

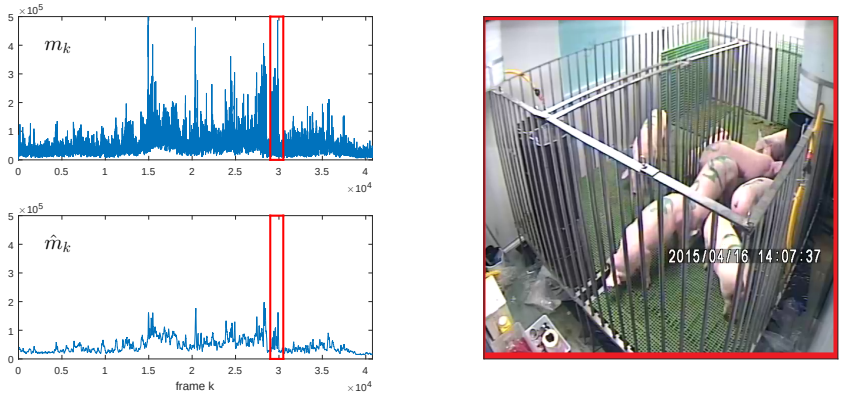


Figure 2.3: Global motion $m_k$ and the corresponding moving average $\hat{m}_k$ for a randomly selected 2h of video footage. The moving average was performed over a window of $K = 45$ frames. The interval of time series boxed in red on the left correspond to the video image on the right, in which the animals were in motion.

### 2.2.5 Motion classification

Three types of motion were observed during the experiment: (1) baseline, when all animals were lying or sleeping, such that motion was minimal or absent; (2) animal motion, when animals were walking, feeding or playing; and (3) human-related motion, when veterinary practices were carried out. Animal movements

sometimes increased in the presence of humans, and also movements of the humans themselves contributed to overall motion when they entered the region of interest being recorded. Therefore we quantified the relative amounts of all three types of motion each day and sought to determine changes in animal motion unrelated to human interaction.

To classify motion as baseline, animal motion or human-related motion, we used short intervals of $S = 15$ seconds (corresponding to $6S = 90$ frames) and we joined consecutive values of $\hat{m}_k$ into $6S$ periods through

$$M_r = \left[\hat{m}_{1+6S(r-1)}, \hat{m}_{6Sr}\right], \forall r \in \{1, 2, ...\}. \tag{2.11}$$

We described each of these periods using five statistical parameters [136]: minimum ($M_r^{\min}$), first quartile ($M_r^{\mathrm{q1}}$), median ($M_r^{\mathrm{med}}$), third quartile ($M_r^{\mathrm{q3}}$) and maximum ($M_r^{\max}$). In addition, we computed the quantity of motion of each $M_r$ as

$$QM_r = \sum_{i=1+6S(r-1)}^{6Sr} \hat{m}_i, \forall r \in \{1, 2, ...\}. \tag{2.12}$$

Firstly, we analysed animal behaviour before experimental infection in order to quantify healthy levels of motion. To do this, we clusterized the short series $M_r$ corresponding to the pre-infection phase by using the $k$-means algorithm (*kmeans* command in Matlab) with the five statistical values ($M_r^{\min}$, $M_r^{\mathrm{q1}}$, $M_r^{\mathrm{med}}$, $M_r^{\mathrm{q3}}$, $M_r^{\max}$), which assigned the input observations into $k$ clusters through step-wise cluster centroid estimation based on optimisation of within- and between-cluster dispersion [91]. To estimate the input number of clusters $k$, we used the gap criterion algorithm (*evalclusters* command in Matlab), which scored the so-called gap statistic associated to a proposed number of clusters based on the minimisation of within-cluster dispersion [147]. Consequently, the maximum value of the gap statistic is associated to the optimal number of clusters (assuming less than 50). Finally, all $k$ clusters (and the $M_r$ records therein) were assigned to baseline (group $G_1$), animal motion (group $G_2$) and (3) human-related motion (group $G_3$).

Secondly, we used a support vector machine (SVM) pairwise classifier to assign $M_r$ records to groups $G_1$, $G_2$ and $G_3$ in the entire experiment. To do this, half the $M_r$ records in the pre-infection phase were randomly assigned to the so-called training dataset and the other half of records to the so-called testing dataset. To train the SVM model, we used the *svmtrain* command in Matlab with the training dataset; then, we evaluated the accuracy of the model using the *svmclassify* command with the testing dataset. This verified model was then used to analyse the remaining phases of the experiment (Figure 2.1).

### 2.2.6 Changes in animal motion as a result of infection

We expected that total daily motion and time spent lying/sleeping and feeding/walking would change after infection with ASF virus. To measure this reliably, we used the Wald-Wolfowitz runs test (*runstest* command in Matlab) at significance levels of 90%, 95% and 99%, corresponding to respective p-values less than 0.1, 0.05 or 0.01, which evaluated whether consecutive daily motion values were randomly distributed around the average value of the considered period. First, runs tests were performed during the pre-infection phase (days 1-11) to verify random distribution of motion values around the average value. Then, runs tests were repeated by expanding the window by one additional day until day 18, when clinical signs were detected. This allowed us to determine whether and when consecutive motion values significantly diverged from the average value, indicating ASF.

## 2.3 Results

### 2.3.1 Motion classification

The gap statistic showed a maximum of 3.6955 for 32 clusters, with the next highest values being 3.6676 for 28 clusters and 3.6653 for 34 clusters. Therefore, we considered $k = 32$ clusters in the $k$-means algorithm. Figure 2.4 shows the centroids of these clusters and the percentage of $M_r$ records classified in each one. After comparing many video sequences associated with $M_r$ records across all clusters, we assigned cluster 1 to baseline ($G_1$), clusters 2-5 to animal motion ($G_2$), and clusters 6-32 to human-related motion ($G_3$). As expected, the lowest motion values corresponded to baseline, which occurred during 70.43% of the pre-infection period. Intermediate motion values corresponded to animal motion, which occurred during 27.41% of the pre-infection period. Nearly one quarter of records (23.56%) fell into cluster 3, when animals were feeding and walking, while 3.28% of records fell into cluster 5, when animals were playing and so were highly active. Finally, human-related motion was greatest occurring during 2.16% of the time into clusters 6-32.

The SVM classifier showed accuracy of 95.66–98.87% at classifying records to groups $G_1$, $G_2$ and $G_3$ (Table 2.1). Conversely, only 1.64% of records in $G_1$ were incorrectly assigned to $G_2$, 1.13% of records in $G_3$ were incorrectly assigned to $G_2$ and records in $G_2$ were incorrectly assigned to $G_1$ (1.14%) or $G_3$ (3.20%).

| group | classified in group | | |
|---|---|---|---|
| | $G_1$ | $G_2$ | $G_3$ |
| $G_1$ | 98.36% | 1.64% | 0.00% |
| $G_2$ | 1.14% | 95.66% | 3.20% |
| $G_3$ | 0.00% | 1.13% | 98.87% |

Table 2.1: Percentages of records in the testing dataset correctly and incorrectly assigned to groups $G_1$, $G_2$ and $G_3$ by the pairwise SVM- classifier.

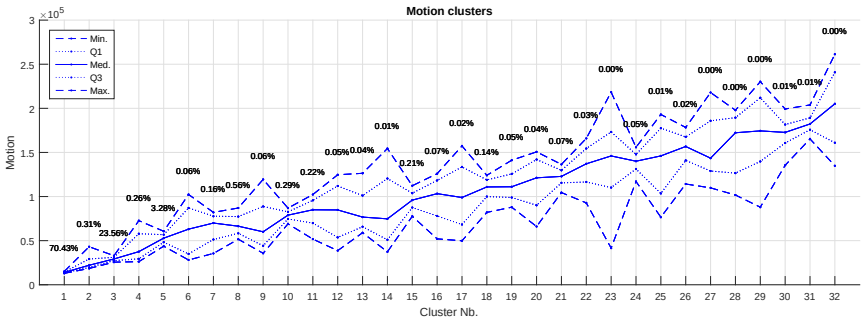


Figure 2.4: Minimum (Min.), first quartile ($Q_1$), median (Med.), third quartile ($Q_3$) and maximum (Max.) centroid values for the $k = 32$ clusters computed through $k$-means algorithm describing motion over 15-second intervals (90 frames) during the pre-infection period. Percentages refer to the total number of records in each cluster.

### 2.3.2 Changes in animal motion as a result of infection

Table 2.2 shows the proportions of time spent each day in baseline, animal motion and human-related motion. After the virus inoculation, lying/sleeping time increased substantially and feeding/walking time decreased proportionally across the following post-infection phases. In fact, in the first four days the animal motion decreased 6.59% and, before the detection through clinical signs, 10.27% respect to the pre-infection phase.

| Type of motion | Pre-infection | Infection | PCR detection | Clinical detection |
|---|---|---|---|---|
| Baseline | 70.43% | 77.91% | 81.65% | 89.18% |
| Animal | 27.41% | 20.82% | 17.14% | 9.36% |
| Human-related. | 2.16% | 1.27% | 1.21% | 1.46% |

Table 2.2: Percentage of time spent each day in baseline, animal motion and human-related motion for each of the four experimental phases defined in Figure 2.1.

Comparison of the percentages of records $M_r$ classified by the SVM approach into groups $G_1$, $G_2$ and $G_3$ for each day of the experiment (Figure 2.5A) shows that human-related motion occurred during only a fraction of the time, but it made a disproportionate contribution to the amount of motion generated, $QM_r$ (Figure 2.4). Therefore, such motion was excluded from the analysis, and the total daily motion (as defined in Equation (2.12)) was re-calculating by accumulating only daily baseline and animal motion (Figure 2.5B). We executed the SVM-classifier for the remaining days of the experiment. Figure 2.5A shows the percentage of records, $M_r$, classified in groups $G_1$, $G_2$ and $G_3$ per day of the experiment; that is, the time classified as baseline motion, animal in motion and human interaction per day. Throughout the experiment, the human interaction represented a small percentage of the records but the quantity of motion generated, $QM_r$, was excessively high (see Figure 2.4) and, therefore, it was discarded for analysis. Figure 2.5B shows the total quantity of motion (see Equation (2.12)) accumulated per day but excluding the

human interaction records.

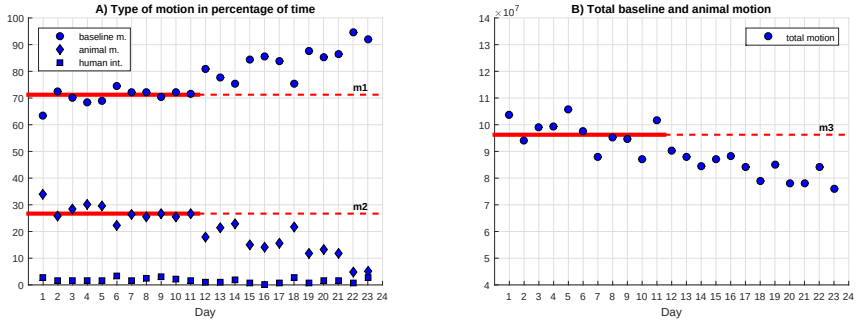


Figure 2.5: A) Percentage of time each day involving baseline, animal motion or human-related motion. Mean values over the first 11 days are shown as red lines for baseline ($m_1$) and animal motion ($m_2$). B) Total daily motion, excluding human-related motion. The mean value for the first 11 days is shown as a red line ($m_3$).

## Pre-infection phase (days 1-11)

Aside from day 1, when animals were slightly hyperactive because of their new environment, they spent fairly constant proportions of each day lying/sleeping and feeding/walking. These proportions varied randomly around the respective mean values of $m_1$ and $m_2$. Similarly, total daily motion varied around the mean value $m_3$.

### Infection phase (days 12-15)

Following experimental infection with ASF virus, the proportion of time spent each day in baseline tended to increase, while the proportion of time spent in animal motion tended to decrease, as did the total daily motion. On days 12 and 15, the percentage of time each day in baseline was higher than on any other previous day of the experiment, while the percentage of time in animal motion was lower than on any other previous day, pointing to a probable effect of ASF. On day 14, total daily motion was lower than on any other previous day.

### PCR detection phase (days 16-18)

The trends observed during the infection phase continued during the PCR detection phase. Although total daily motion on days 15-16 was similar to that during the pre-infection phase, percentages of baseline and animal motion remained quite low. On day 18, when the first clinical signs of ASF were detected, animal motion increased considerably, but not total daily motion. It was possible that animal motion was higher on this day because of residual influence from the presence of

workers, yet total motion showed low values due to the deceleration of movements recorded.

### Clinical detection phase (days 19-23)

The trends observed during the infection phase and PCR detection phase worsened during the clinical detection phase, with the percentage of baseline increasing and percentage of animal motion decreasing. This worsening should be interpreted with caution because the animals began to die rapidly from day 18 onwards. Therefore the data collected during this phase is unlikely to be useful for assessing the ability of our approach to detect ASF early.

### Runs test

Runs tests were carried out for progressively longer time windows during the 23-day recording period in order to examine on what day the changes in baseline, animal motion and total daily motion deviated significantly from pre-infection values (Table 2.3). These runs tests focused on days 11-18, the period critical for early ASF detection (Figure 2.5). Runs tests for the period from days 1 to 11 suggested that, as expected, percentages of baseline and animal motion as well as total daily motion varied randomly around their mean values. Runs tests for longer periods indicated that although the percentage of baseline increased after inoculation and the percentage of animal motion and amount of total daily motion decreased, these changes did not deviate significantly from pre-infection values until day 15, when the differences became significant at the 90% level. The differences were significant at the 95% level on day 16 and at the 99% level on day 18.

| Runs | Day interval | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| Baseline | [1,11] | [1,12] | [1,13] | [1,14] | [1,15] | [1,16] | [1,17] | [1,18] |
| $\|Z\|$ | 0.03 | 0.21 | 0.02 | 0.83 | 1.93 | 2.22 | 2.45 | 2.65 |
| Sig. (2-tailed) | 0.97 | 0.84 | 0.98 | 0.40 | 0.05* | 0.03** | 0.01** | 0.01*** |
| Animal motion | [1,11] | [1,12] | [1,13] | [1,14] | [1,15] | [1,16] | [1,17] | [1,18] |
| $\|Z\|$ | 1.10 | 0.21 | 1.19 | 0.83 | 1.92 | 2.21 | 2.45 | 2.65 |
| Sig. (2-tailed) | 0.27 | 0.83 | 0.23 | 0.40 | 0.05* | 0.03** | 0.01** | 0.01*** |
| Daily motion | [1,11] | [1,12] | [1,13] | [1,14] | [1,15] | [1,16] | [1,17] | [1,18] |
| $\|Z\|$ | 0.61 | 0.30 | 0.56 | 1.25 | 1.70 | 1.88 | 2.04 | 2.19 |
| Sig. (2-tailed) | 0.54 | 0.76 | 0.57 | 0.21 | 0.09* | 0.06* | 0.04** | 0.03** |

*, ** and ***: significance at 90%, 95% and 99%, respectively.

Table 2.3: The $|Z|$ statistic and 2-tailed significance (Sig.) for runs tests of percentages of daily time spent in baseline or animal motion (Figure 2.5A), or of total daily motion (Figure 2.5B). Runs tests were performed on progressively longer intervals of the experimental period, from days 1-11 to days 1-18.

## 2.4 Conclusions and discussion

Video processing of livestock offers the possibility of non-intrusive animal monitoring in real time. Here we aimed to examine whether such monitoring could reliably detect decreases in pig motion in the early stages of ASF infection. Our results in this pilot study showed that, indeed, we detected a significant decrease in motion at just four days after experimental infection with the ASF virus. This was the same day that the virus was detected in blood using qPCR, and a full three days before clinical signs of ASF were observed.

The proposed method comprises motion estimation, clustering/classification of motion patterns and recognition of these patterns via supervised learning. "Optical flow" is a widely used algorithm to measure the apparent velocities of objects in video streams that allowed us to quantify motion between consecutive frames throughout the entire experiment. We divided the total recording period into intervals of a few seconds and classified motion in each interval into (1) baseline (little or no motion), (2) animal motion and (3) human-related motion. An SVM pairwise classifier was used to train the model to recognise the three types of motion based on a randomly selected 50% of the pre-infection dataset, and this ability was then assessed using the other half of the same dataset. After establishing an accuracy >95%, we applied the classifier to the data collection during the post-infection phases of the experiment. This allowed us to detect an increase in the proportion of baseline activity and concomitant decreases in animal motion and total daily motion. To assess the possibility that these results might be due only to chance, we performed runs tests on data over progressively longer time windows. This analysis indicated that the changes in motion attributable to ASF became significant at four days after infection (90% level).

In our study, a slight perturbation in the global motion was recorded, even when no animals were in the pen (before the start of the experiment) or when the animals did not move, probably reflecting the SD resolution of the camera. Our efforts to treat these perturbed time series as autoregresive models (ARMA/ARIMA) were unsatisfactory, so we treated them as normally distributed and as contributing a constant amount of daily motion to the observed total. It is possible that alternative approaches to noise reduction may allow more refined classification and recognition of motion types, such as differentiation between feeding and walking. In an ideal case, such approaches should be adaptable to a range of farm and animal types through machine learning algorithms, as used here. On the other hand, the runs tests used here were able to detect changes in 11 days but, considering real implementation in farms with longer periods without of no infection, it could be more effective the monitoring through quality control techniques.

Likely important to the success of our approach is filtering out animal motion largely or entirely induced by veterinarians and technicians visiting the pen, as well as movement of these workers mistakenly recorded as animal motion during the experiment. While in practice only a fraction of daily motion in our study was human-related, this motion occasionally produced false spikes and so needed to

be removed. Future work should examine the possibility of extending Optical flow through the implementation of animal tracking and recognition algorithms [138, 152] in order to eradicate definitely human-related motion and to enhance the individual animal monitoring. Therefore, our results with an SD resolution camera suggest that future work should focus on high definition (HD) cameras, which can automatically adjust contrast and brightness and so which may further reduce background noise and blurring.

Our results illustrate the promise of video processing for early detection of ASF. Many other livestock diseases might present slowdown in motion caused by fever in animals, such as foot-and-mouth disease, classical swine fever, Rift Valley fever, African horse sickness or BT. The possibility to extend the technology presented here to detect rapidly other diseases might improve the current surveillance protocols substantially. In fact, we showed that four days after inoculation, the system launched an alert of ASF-suspicion that, under real farm conditions, probably could be interpreted for the veterinary services to take blood test and confirm an outbreak three days before the first clinical signs.

# Chapter 3

# Epidemic modeling and simulation

## Articles in peer reviewed journals

- **E. Fernández-Carrión**, B. Ivorra, B. Martínez-López, A.M. Ramos, J.M. Sánchez-Vizcaíno. "*Implementation and validation of an economic module in the Be-FAST model to predict costs generated by livestock disease epidemics: Application to classical swine fever epidemics in Spain*". Preventive Veterinary Medicine 126, pp. 66-73 (2016).
- B. Martínez-López, B. Ivorra, **E. Fernández-Carrión**, A.M. Pérez, A. Mendel-Herrero, F. Sánchez-Vizcaíno, C. Gortázar, A.M. Ramos, J.M. Sánchez-Vizcaíno. "*A multi-analysis approach for space-time and economic evaluation of risks related with livestock diseases: The example of FMD in Peru*". Preventive Veterinary Medicine 114, pp. 47-63 (2014).
- B. Martínez-López, B. Ivorra, A.M. Ramos, **E. Fernández-Carrión**, T. Alexandrov, J.M. Sánchez-Vizcaíno. "*Evaluation of the risk of classical swine fever (CSF) spread from backyard pigs to other domestic pigs by using the spatial stochastic spread model Be-FAST: The example of Bulgaria*". Veterinary Microbiology 165, pp. 79-85 (2013).

## Proceedings

- A.M. Ramos, B. Ivorra, **E. Fernández-Carrión**, B. Martínez-López, D. Ngom, J.M. Sánchez-Vizcaíno. "*Be-CoDIS and Be-FAST: Mathematical models to predict the spread of human and livestock diseases with real data. Application to the 2014-15 Ebola Virus Disease epidemic and livestock diseases*". VI International Conference on Environmental, Industrial and Applied Microbiology, BioMicroWorld2015 (23 February 2016), Madrid (Spain). Oral presentation.

- **E. Fernández-Carrión**, B. Ivorra, B. Martínez-López, A.M. Ramos, J.M. Sánchez-Vizcaíno. "*Implementation and validation of an economic module for the epidemiological model Be-FAST to predict the costs generated by livestock diseases epidemics*". 14 International Symposia on Veterinary Epidemiology and Economics, ISVEE (3–7 November 2015), Merida (Mexico). Poster.

- **E. Fernández-Carrión**. "*Be-FAST: a time-spatial stochastic spread model for studying the transmission of infectious livestock diseases and its economic impact*". II Workshop on Modelling and Simulation of Epidemics (26 May 2015), Madrid (Spain). Workshop.

- M.A. Alkhamis, B. Martínez López, B. Ivorra, **E. Fernández-Carrión**, L. Mur, T. Alexandrov, A.M. Ramos, J.M. Sánchez-Vizcaíno, 2014. "*Evaluation of the magnitude and duration of potential African Swine Fever (ASF) epidemics in Bulgaria using a spatial and stochastic disease spread model (Be-FAST)*". Annual Conference of the Society for Veterinary Epidemiology and Preventive Medicine (26–28 March 2014), Dublin (Ireland). Poster.

- B. Ivorra, **E. Fernández-Carrión**, B. Martínez-López, J.M. Sánchez-Vizcaíno, A.M. Ramos. "*Be-FAST – Between Farm Animal Spatial Transmission: An epidemiological model for studying the spread and the economic impact of animal diseases*". Annual Meeting of the Society for Veterinary Epidemiology and Preventive Medicine, SVEPM (23 February 2013), Madrid (Spain). Poster.

## Copyrighted software

- **E. Fernández-Carrión**, B. Ivorra, B. Martínez-López, A.M. Ramos, J.M. Sánchez-Vizcaíno. "*Be-FAST*". Request number, M-2614-15. Submission date, 17 April 2015.

## 3.1 Introduction

The spread of virus from infected animals to susceptible ones may occur either by direct or indirect contacts during livestock diseases epidemics. Direct transmission implies animal-to-animal effective contact from an infectious animal to a susceptible one. Indirect transmission requires an effective contact between a contaminated fomite (i.e. vehicle, vectors, material or people) and a susceptible animal. Generally, the main routes of transmission (spread patterns) as well as the sanitary and economic consequences (magnitude) associated to each epidemic depend on the specific features of the disease, the demographic distribution of the host population and the timing/effectiveness of the applied control measures applied [98]. These aspects imply that any work based on assessing spread patterns and magnitude of possible epidemics implies a degree of uncertainty caused by the virulence of the strain, the livestock population dynamics, the local socio-cultural and ecological factors, or even external factors that may vary the effectiveness of the control measures. Epidemic modeling focuses on the development of decision-making tools aimed to reduce this uncertainty. In this regard, epidemic simulation tools allow to reproduce or extrapolate known epidemics to different target regions or periods based on expert assessment and historical information for evaluating the risks associated to the spread patterns and magnitude.

In the literature, there are several models designed to simulate the potential spread of livestock diseases into specific regions [130, 72, 139, 108, 70, 143]. Generally, these models are based on Monte Carlo simulations of stochastic variables that generate different output epidemics (scenarios). The outcomes of all scenarios provide quantitative estimates on the magnitude and duration of potential epidemics that help to identify the main risk factors for the diseases spread, to evaluate the efficiency of control measures and to identify specific diffusion patterns. It is also remarkable that such methods enable to approach a first evaluation of potential spread in disease-free areas or areas where livestock farming has changed considerably.

Be-FAST (*Between and within Farm Animal Spatial Transmission*) is a computer program based on a time-spatial stochastic spread mathematical model for studying the transmission of infectious livestock diseases. This model has been described in Martínez-López et al. (2011) and strictly validated in Martínez-López et al. (2012) [98, 99]. In contrast to other epidemic modeling tools, Be-FAST offers novel approaches by considering specific farm-to-farm contact network in the study region, which implies substantial improvements in spatial analysis, and within farms spread, which is commonly omitted or poorly described in other models. In a few words, Be-FAST simulates the within-farm transmission using a modified Susceptible-Infected (SI) compartmental model meanwhile the between-farm transmission is assumed to occur by direct contacts (i.e. animal movements) and indirect contacts (i.e. local spread, vehicle and person contacts) considering the spatial location of farms. Furthermore, the model implements control measures ruled by European legislation (i.e. depopulation of infected farms, movement restriction, zoning, surveillance, contact tracing).

The first versions of Be-FAST focused on the simulation of classical swine fever (CSF) epidemics in the region of Segovia (Spain). Available input data sets are crucial in order to obtain accurate outcomes. In this case, regional government of Castilla-León provided detailed input data of live animal movements and farm locations. This region was severely affected by two historical CSF epidemics (in 1997-98 and 2001-02) becoming a good scenario for model validation. This work has been widely described in Martínez-López et al. (2011) [98]. The Be-FAST model aims to be an accurate decision-support tool adaptable to several notifiable diseases in different livestock frameworks (under similar local legislation for control and eradication), useful in the development of prevention and control strategies. The algorithm has sufficient features to reproduce epidemics of different notifiable diseases.

This chapter does not intend to describe the Be-FAST algorithm. Such information can be obtained in Ivorra et al. (2013) [68]. Nevertheless, we start in Section 3.2 with a brief explanation of this model and main concepts in epidemiological modeling of the Be-FAST algorithm, necessary to understand the following sections. Here, we aim to adapt epidemiological features of other notifiable diseases, regions and periods in Be-FAST for risk analysis and to implement additional processes for studying the economic impact of epidemics.

In Section 3.3.1, we aim to explore the potential spread of CSF from backyard pigs to other domestic pigs by using Be-FAST. In Bulgaria, backyard pigs have been recognized as key players for disease occurrence, not only for CSF [119, 92]; however, there are no studies addressing and quantifying their epidemiological role in the potential CSF transmission to other domestic pigs. This lack of studies is, most likely, associated with the scarceness of complete and reliable information about backyard pig demographics and contact patterns, which are key factors to estimate disease transmission in countries/regions where backyard pig production is predominant. Specifically, Bulgaria holds a large number of backyard farms (96% of the total farms). A better understanding of the role that backyard pigs have in CSF transmission will enhance the CSF eradication program, supporting the implementation of better surveillance and control strategies. Methods and results presented in Section 3.3.1 may be useful to guide risk-based interventions not only in Bulgaria, but also in other similar countries where backyard pig production is predominant.

In Section 3.3.2, we aim to evaluate the spatial spread of foot-and-mouth disease (FMD) in mixed farms (i.e. farms with more than one specie). FMD have has had a severe impact on the livestock industry, limiting the trade of animals and animal products [111, 127]. For this reason many countries, in collaboration with international organizations have established regional or national projects to progressively control and finally eradicate FMD. Peru has currently achieved *FMD-free without vaccination* status in more than 98.3% of the country and the remaining 1.7% is *FMD-free with vaccination* [25]. Livestock production in Peru represents 43% of the agricultural GDP (gross domestic product) and is mostly extensive or semi-extensive [66]. It has been estimated that pastures in high Andes sustain the 78.8% of cattle, 96.2% sheep, 100% of South American camelids [115]. Production is mainly

for self-consumption, representing the economic basis of many rural communities, which determines the predominant small-medium farm sizes and the large number of mixed farms.

Any future ban on FMD vaccination associated with eradication will bring a certain risk for Peru due to the potential for FMD re-introduction into naïve populations. Consequently, surveillance and control programs will need to be re-designed to account for this new epidemiological case (i.e. freedom from disease) with the ultimate goal of designing and implementing the most cost-effective measures to preserve the FMD-free status. Thus, it is crucial to obtain detailed information and data on risk metrics associated with alternative epidemiological scenarios and measures of uncertainty and variability in order to design risk-based surveillance systems. Such output cannot be obtained just analyzing historical outbreaks and requires the integration of a variety of tools, borrowing concepts and methods from diverse disciplines, such as spatial risk analysis.

Finally, in Section 3.4, we extend the capabilities of the model to economic assessment. The large economic investments that the countries apply to fight livestock epidemics (for instance, by implementing control measures) generate an important demand for economic decision making tools. In particular, CSF is a highly contagious viral disease that affect wild and domestic swine. It is considered one of the most economically damaging diseases in the swine industry worldwide [16, 61]. The pig sector in the EU maintains a high level of production and exports, contributing more than €32 billion per year to the EU economy [123], the eradication of CSF has become a chief priority in the EU. In fact, the EU funding for the eradication of CSF from 2005 to 2009 amounted to €17 million [123]. This financial effort has resulted in a significant reduction of the CSF outbreaks in most of the countries, with the practical eradication of the disease in domestic pig populations. Nevertheless, sporadic CSF outbreaks still occur in many European countries.

Spain is the second largest producer of swine in the EU. It was declared CSF-free in 1988, but two CSF incursions have occurred since then: one in 1997-98 that affected the provinces of Lleida, Seville, Segovia and Saragossa; and another in 2001-02 that affected the provinces of Lleida, Castellón, Valencia, Cuenca and Barcelona. Both epidemics caused significant estimated economic costs (Table 3.1). Although the characteristics and economic consequences of each CSF epidemic depend on the location, time period and type of holdings infected, strict control measures are always required in order to stop the spread of disease and eradicate it, mandated by Spanish and EU regulations [122].

| year | depop. farms | months | culled animals | economic losses |
|---|---|---|---|---|
| 1997-98 | 99 | 16 | 609,147 | €60 million |
| 2001-02 | 48 | 11 | 378,407 | €48 million |

Table 3.1: Summary data of CSF epidemics in Spain since 1988 [120]. Number of depopulated farms, duration of the epidemic in months, number of culled animals and estimated costs to compensate farmers by culled animals.

All the control measures generate economic costs that are supported by government authorities and the swine industry. The study of the potential spread patterns of CSF into an area may help to identify risk zones to improve the prevention and management of future outbreaks. The main goals are to develop an economic model to evaluate various economic costs of simulated livestock disease outbreaks, to include this model as a module in Be-FAST and to validate it using historical data on the two most recent CSF incursions in Spain. The value of this study is to provide a useful tool for authorities and insurance companies to estimate an initial budget needed to fight against a particular disease in a specific area, as well as predict how costs will evolve during an epidemic. Finally, we use our model to analyse relationships among model behaviour, public health and economic impact. We simulate different scenarios in order to analyse how each type of cost evolved over the course of the epidemic. We compare our model output with real historical economic data for CSF epidemics in Spain.

## 3.2 Be-FAST epidemic simulating tool

### Compartmental disease modeling

The interest in infectious disease modeling begun around XIX century and the most representative progress in this field was shown in the early twentieth century, with Hamer (1906), Ross (1908-1911) and Kermack-McKendrick (1927) works. More specifically, Hamer hypothesized that incidence of infection is proportional to the product of the densities of susceptible and infected individuals cite. This phenomenon is called the *mass-action* mixing assumption, and is a basic concept of epidemic modeling [13]. Later, Ross conducted a study on malaria transmission between mosquitoes and humans, concluding that controlling the mosquito population could avoid cyclic outbreaks of malaria. In fact, Ross formulated a mathematical model on a vector-borne disease able to estimate critical thresholds to avoid endemic spreads [17]. Kermack and McKendrick work set the first ordinary differential equations which described the fluctuation of the susceptible population ($S$), the infected population ($I$) and the recovered population with immunity ($R$) over time in an epidemic [75]. These equations were the first to describe the so-called $SIR$ model which is used as the basis of many current epidemic models, called compartmental models. Among those models, the $SEIR$ model is quite similar to the $SIR$ model but it incorporates an additional exposed population ($E$) associated to infected individuals that are not yet infective during a period of time (called incubation period). The $SEIR$ model fits well the livestock diseases such as CSF or FMD. Here, we only focus on a brief description of key issues of compartmental models in order to have better understanding of the Be-FAST model. For more detailed information, see, for instance, Diekmann et al. [35] and Brauer et al. [18].

The compartmental models start off with classifying the host population ($N$) into different subsets depending on the infectious status of each individual. The basic compartmental $SIR$ model divide the population into: susceptible population

($S$), individuals who have no immunity to the infectious agent; infectious population ($I$), individuals who are currently infected and can transmit the infection to susceptible individuals; and removed population ($R$), individuals who are immune to the infection, and consequently do not affect the transmission dynamics in any way [38]. In a naive population, every individual is susceptible to be infected. Once a newly infected individual is introduced, he can transmit the disease to susceptible persons during its infectious period of the disease; that is, the period for which an infected person can transmit a pathogen to a susceptible host. This implies an increment of the infected population and the corresponding decrease of the susceptible population. After the infectious period, the infected individuals become immune to the infection, increasing the recovered population. The changes in the population of all these compartments depend on the specific epidemiological characteristics of the disease. The goal of the $SIR$ model is to study the variation in population of all those compartments over time.

In compartmental models, the transitions between the compartments of the population are assumed to be governed by rates. Specifically, in the basic $SIR$ model (Figure 3.1), assuming a constant population $N = S + I + R$, the model is governed by [75]

$$\begin{cases} \dfrac{dS}{dt} = -\beta \dfrac{SI}{N} \\ \dfrac{dI}{dt} = \beta \dfrac{SI}{N} - \gamma I \\ \dfrac{dR}{dt} = \gamma I. \end{cases} \tag{3.1}$$

In this System, $\gamma$ denotes *the recovery rate* of persons in state $I$ to person in state $R$ (i.e. infectious persons leave the infective status at rate $\gamma I$ per unit time). Moreover, $\beta$ denotes the *effective contact rate per individual* which measures the number of contacts between infected and susceptible persons sufficient to transmit the disease per unit time. Therefore, *susceptible persons* leave the susceptible status with a *transition rate* $\beta I$ multiplied by the proportion of susceptible individuals $S/N$ per unit time. Thus, the length of the infective period assumes an exponential distribution with mean $1/\gamma$, called *mean infectious period* [17].

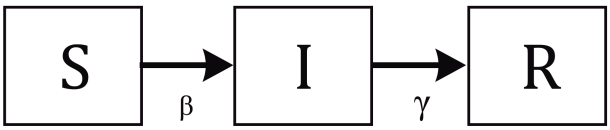


Figure 3.1: $SIR$ model flowchart.

Similarly, the $SEIR$ model takes in consideration the exposed compartment ($E$) situated between susceptible and infective persons (Figure 3.2). Persons in state $E$ are recently infected and cannot transmit the disease due to specific features of some infectious diseases (called latent period). In these cases, the time to pass from state

$E$ to state $I$ is governed by a *mean exposed period* $1/\delta$. The $SEIR$ model can be written as

$$\begin{cases} \dfrac{dS}{dt} = -\beta \dfrac{SI}{N} \\ \dfrac{dE}{dt} = \beta \dfrac{SI}{N} - \delta E \\ \dfrac{dI}{dt} = \delta E - \gamma I \\ \dfrac{dR}{dt} = \gamma I. \end{cases} \tag{3.2}$$

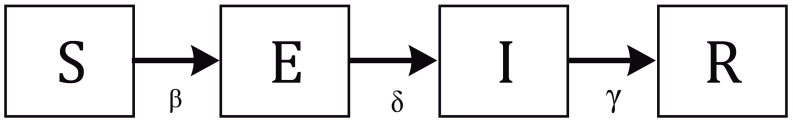


Figure 3.2: $SEIR$ model flowchart.

### General description of Be-FAST

Be-FAST algorithm is based on a Monte Carlo method that generates $M \in \mathbb{R}$ scenarios of possible epidemic evolutions. The algorithm is summarized in Figure 3.3. Input data consist of a user-specified real database of farms and of movements between farms, based on animal trading. The algorithm assumes that at time $t = 0$, every farm is free of disease and therefore in a susceptible status except for a predefined number of randomly selected farms (called *index cases*) that are assumed to contain a given number of infected animals. During a period of time $[0, T]$, with $T \in \mathbb{N}$ being the maximum number of days in the simulation, the disease spreads *within-farm* through a modified Susceptible-Infected model with animals playing the role as individuals, and *between-farm* through an Individual-Based model with farms playing the role of individuals. Every day of the simulation, the authorities may detect contaminated farms, in which case the applicable control measures are activated in order to stop the spread, and consequently the epidemic. Afterward, the simulation concludes and the next one begins until the last scenario $M$.

### Within farm transmission

Be-FAST model classifies the host animal population into different compartments. Specifically, when an animal is not infected, it is categorized in the Susceptible ($S$) status. Once it is infected, it passes successively through the following states [39]:

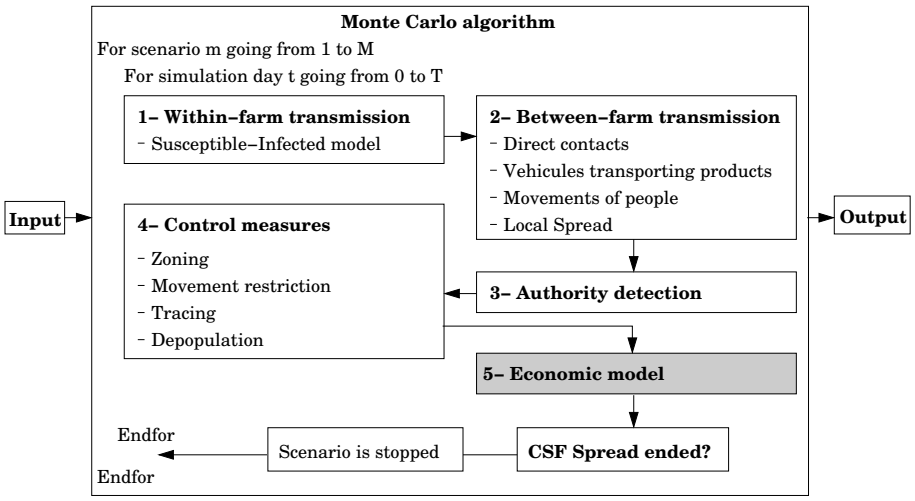


Figure 3.3: Main structure of the Be-FAST algorithm.

- Infected ($E$): The animal is infected by the virus but cannot infect other susceptible animals and has no visible clinical signs.

- Infectious ($I$): The animal has been infected by the virus, can infect to other susceptible animals but has no visible clinical signs.

- Clinical signs ($C$): The animal develop visible clinical signs associated to the disease and still may infect other animals.

We note that Be-FAST model is based on a modified version of the $SEIR$ model for which the Recovered ($R$) status has been removed due to the fact that the culling of infected animals is usually accomplished before the swines reach this state. Nevertheless, we consider a Clinical signs ($C$) status important for disease detection and to trigger the control measures for eradication (see Section 3.2). The time interval between each status is governed by the following periods:

- Latent period: Time duration from the beginning of the Infected status till the animal can spread the disease. After such period, the animal jumps to Infectious status.

- Incubation period: Time duration from the beginning of the Infectious status of an animal to the development of clinical signs. After such period, the animal jumps to Clinical signs status.

Analogously, farms are classified one of the following status [72]: Susceptible ($S$), if all animals in the farm are susceptible; Infected ($E$), if at least one animal in the farm is infected; Infectious ($I$), if at least one animal in the farm is infectious; and Clinical signs ($C$), if at least one animal in the farm shows clinical signs.

From an implementation point of view, the disease spread within a farm $i \in \{1, ..., NF\}$ (where $NF$ as the number of farms) was designed according to a modified SI based on the following equations:

$$\begin{cases} \Delta S_i(t + \Delta t) = S_i(t) - \beta \dfrac{S_i(t) I_i(t)}{N_i(t)} \Delta t, \\ \Delta I_i(t + \Delta t) = I_i(t) + \beta \dfrac{S_i(t) I_i(t)}{N_i(t)} \Delta t, \end{cases} \tag{3.3}$$

considering the total animal population at time $t$ in a farm $i$ as $N_i(t) = S_i(t) + I_i(t)$, daily dynamic ($\Delta t = 1$ day as minimum value) and $S_i(t), I_i(t) \in \mathbb{N}, \forall t \in [0, T]$. The evolution of the proportion of infected animals in a farm $i$ is governed by equations (3.3) but once one animal is infected in this farm, the model estimates randomly the latent and incubation periods according to the disease characteristics in order to pass from Exposed status to Infected status and from Infected status to Clinical Signs status.

### Between farms transmission

The farm-to-farm transmission of the disease was modeled by considering both direct and indirect contacts. Transmission of direct contacts from an infected farm $i$ to a susceptible farm $j$ at day $t$ is based on the network of animal shipments, given as input data set. Although this database usually contains previous years' information, the shipments may be assumed similar from one year to another with slight variations. Thus, Be-FAST uses random movements, generated from the pig transports data, instead of using the exact ones (using exact or random shipments does not impact dramatically the model outputs). Due to the construction process described below, these random movements exhibit similar characteristics (i.e. farms of origin and destination, date and number of moved pigs) than the real ones. However, this process allows to consider, with a low probability, transports to or from farms not included in the shipment database (for instance, due to errors). More precisely, the model computes the number of movements occurring by considering a Poisson distribution (for the case of CSF) with mean being the number of all movements occurring at day $t$ in the input database. Then, for each simulated movement [68]:

- We select randomly the farm of origin of the movement $i$ and the farm of destination of the movement $j$ ($i, j \in \{1, ..., NF\}$ and $i \neq j$) by considering the discrete probability $\mathbb{P}_M((i, j) = (k, l)) = M_{\text{mov}}(k, l) / \big( \sum_{m=1}^{NF} \sum_{n=1, n \neq m}^{NF} M_{\text{mov}}(m, n) \big)$, where $k, l \in \{1, ..., NF\}$, $k \neq l$ and $M_{\text{mov}}(k, l) \in \mathbb{R}$ is the number of movements from farm $k$ to farm $l$ in the database.
- We compute $w_{(i,j)} = \min \big\{ \text{Ceil}\big(\overline{w_{(i,j)}}(N_i(t)/N_i(0))\big), N_i(t) \big\}$, the number of pigs shipped during this movement from farm $i$ to farm $j$. In the previous expression, $\overline{w_{(i,j)}} \in \mathbb{R}$ denotes the mean number of pigs moved when considering all the shipments between farms $i$ and $j$ in our database and $\text{Ceil}(x)$ returns the nearest integer greater than or equal to $x \in \mathbb{R}$.

- Next, we move $w_{(i,j)}$ pigs from the origin farm $i$ to the destination farm $j$. These pigs are selected randomly in $S_i(t)$ and $I_i(t)$, considering a discrete uniform distribution. We denote by $w^S_{(i,j)} \in \mathbb{N}$ and $w^I_{(i,j)} \in \mathbb{N}$ the number of susceptible and infected pigs moved in shipment, respectively. Thus, the evolution of pigs in farm $i$ and $j$ is also governed by $S_i(t) = S_i(t) - w^S_{(i,j)}$, $I_i(t) = I_i(t) - w^I_{(i,j)}$, $S_j(t) = S_j(t) + w^S_{(i,j)}$ and $I_j(t) = I_j(t) + w^I_{(i,j)}$, updating Equations (3.3).

- Finally, if $w^I_{(i,j)} > 0$, the state of farm $j$ is set to the state of farm $i$ in one of the following cases: The state of farm $j$ is susceptible ($S$); the state of farm $j$ is infected ($I$) and the state of farm $i$ is infectious ($E$) or clinical signs ($C$); the state of farm $j$ is $E$ and the state of farm $i$ is $C$. In other cases, the state of farm $j$ remains unchanged.

Regarding indirect transmission, we consider the following risks:

**A- Movements of vehicles transporting pigs:**

If the farm of origin of the transport is either in the state $E$ or $C$, the truck transporting pigs is considered as contaminated and, thus, can infect the farm of destination. In that case, we assume that the probability of CSF virus infection in the farm of destination due to the contact with the contaminated vehicle is modeled by using a Bernoulli distribution.

**B- Movements of vehicles and persons:**

The CSF virus spread by contact with vehicles or persons visiting farms is assumed to occur only among the farms belonging to the following assumptions:

- The daily number of contacts with vehicles or person per farm is assumed to be Poisson distributed.

- A vehicle or person can visit a maximum number of farms per day.

- A vehicle or person is contaminated if, previously, it has visited a spreading farm (i.e. a farm either in the state $E$ or $C$).

- The probability of CSF virus infection in a farm per contact with a contaminated vehicle (contaminated person, respectively) is modeled by using a Bernoulli distribution.

Thus, for each simulation day, we build the routes of the vehicles and we simulate the way they spread CSF virus by considering the process described below:

- For each farm, we compute the number of vehicles visiting it by using a Poisson distribution.

- Then, we list the farms that are visited by the vehicles and we rearrange this list, denoted by $L$, randomly (taking into account that a farm cannot be visited twice consecutively).

- Next, a first vehicle visits the first four farms in $L$, following the list order. Each fourth farm, until the end of $L$, we consider a new vehicle (non-contaminated) starting from the next farm in $L$.

- During each simulated trip, a vehicle becomes contaminated at the moment it visits a spreading farm and can infect other farms by considering a Bernoulli distribution.

The same method is used to model the itineraries and the CSF virus spread of persons, but with the corresponding parameters values.

**C- Local spread:**

We assume that the local spread occurs to farms in the proximity of a farm either in the state $E$ or $C$. It is mainly due to the airborne spread and contacts with contaminated neighborhood persons or fomites.

For CSF, the daily probability of infection in farm $j$, due to the local spread from spreading farm $i$ at simulation day $t$, is modeled by considering a Bernoulli distribution with mean $P_{LS}(d(i,j)) \times \big(I_i(t)/N_I(0)\big)$, where $d(i,j)$ is the distance (in meters) between the centroid of farms $i$ and $j$, and $P_{LS}(x) \in [0,1]$ is the mean daily probability of infection due to local spread between two farms at a distance $x$, interpolated from Table 3.2.

| Distance $x$ (meters) | 0 | 150 | 250 | 500 | 1,000 | 2,000 |
|---|---|---|---|---|---|---|
| Value of $P_{LS}(x)$ | 0.02 | 0.014 | 0.009 | 0.0038 | 0.0019 | 0 |

Table 3.2: Interpolation points used to compute the probability of infection due to local spread, $P_{LS}$, per infected animal in origin farm.

**D- New infection and state transition:**

For each new CSFV infection occurring at day $t$ in farm $i$ during the processes described in Paragraphs $A$ to $C$, if $S_i(t) \geq 1$, we infect one new pig in farm $i$ by considering $S_i(t) = S_i(t) - 1$ and $I_i(t) = I_i(t) + 1$, updating Equations (3.3). Furthermore, if the state of farm $i$ is $S$, we change it to $I$.

### Disease detection

A contaminated farm is generally detected by the observation of clinical signs of at least one animal. Thus, the daily probability of detection by Authorities depends on whether the disease has not been previously detected in the region (that is, the index case has not been detected) or whether the local Authorities have already reported an outbreak. In both cases, the probability of detection of clinical signs is based on the specific disease features and local information.

### Control measures for disease eradication

Measures applied by the local Authorities to detect and control target diseases are defined in the global legislation mainly assessed by the OIE [109]. Generally, before the first detection of an infected or infectious farm, the detection occurs once the animal presents clinical signs associated to the disease and the corresponded farmer or veterinarian notifies the suspicion of the infection to the Authorities following a strict protocol [81]. In order to control the potential disease epidemic, the following control measures usually are carried out and also implemented in Be-FAST model:

1. Zoning: Establishing a control and surveillance zone around the infected farms during a limited period of time in order to accelerate the detection of infected farms and movements in the vicinity that may spread disease.

2. Movement restrictions: Outgoing and incoming movements within the considered protection zones are limited during a specified period of time, including movements of animals, vehicles and people.

3. Depopulation: All animals of a detected farm are immediately culled. The premise is cleaned and disinfected. Also supplies and material are also destructed.

4. Quarantine: During a period of time, the farm is declared in quarantine and no animals are allowed in the farm. Furthermore, the farm is disinfected periodically.

5. Tracing: Tracing activities involve the process of determining contacts that have left or entered a detected farm during a time interval preceding the detection. The objective of tracing is to identify potential infectious contacts which may have introduced the virus into the farm.

### List of parameters

Finally, the Table 3.3 summarizes the list of input parameters used for modeling the CSF case in Segovia.

| Parameter | Value/Distribution | Ref. |
|---|---|---|
| Within-farm transmission parameter for farrowing pig farms (1) | $\beta_1 = 8.52$ | [80] |
| fattening pig farms (2) | $\beta_2 = 1.85$ | [80] |
| and farrow-to-finish pig farms (3) | $\beta_3 = 5.18$ | [80] |
| Latent period (days) | Po(7) | [72] |
| Incubation period (days) | Po(21) | [72] |
| Probability of infection by infectious vehicles transporting animals | Be(0.011) | [141] |
| Probability of infection by infectious vehicles transporting products | Be(0.0068) | [141] |
| Probability of infection by infectious people | Be(0.0065) | [141] |
| Radius of control zone (km) | 3 | exo |
| Radius of surveillance zone (km) | 10 | exo |
| Duration of control zone (days) | 30 | exo |
| Duration of surveillance zone (days) | 40 | exo |
| Number of daily contacts with vehicles transporting products per farm | Po(0.4) | [72] |
| Number of daily contacts with people per farm | Po(0.3) | [72] |
| Maximum number of daily farms visited by a vehicle transporting products | 4 | exo |
| Maximum number of daily farms visited by a person | 3 | exo |
| Daily probability of detection of the index case due to clinical signs | Be(0.03) | [72] |
| Daily probability of detection due to clinical signs | Be(0.06) | [72] |
| Probability of detection in the control zone due to clinical signs | Be$(0.98 \cdot I(t)/N(t))$ | exo |
| Probability of detection in the surveillance zone due to clinical signs | Be$(0.95 \cdot I(t)/N(t))$ | exo |
| Probability of detection due to serological test | Be(0.95) | [121] |
| Probability of restricting an animal movement on detected farms | Be(0.99) | exo |
| Probability of restricting a vehicle movement on detected farms | Be(0.95) | exo |
| Probability of restricting a people movement on detected farms | Be(0.80) | exo |
| Probability of restricting an animal movement within the control zone | Be(0.95) | exo |
| Probability of restricting a vehicle movement within the surveillance zone | Be(0.90) | exo |
| Probability of restricting a people movement within the control zone | Be(0.70) | exo |
| Probability of restricting a people movement within the surveillance zone | Be(0.70) | exo |
| Probability of restricting movements outside the protection zones | Be(0.40) | exo |
| Duration of general movement restriction (days) | 30 | exo |
| Maximum number of daily depopulated farms | Po(20) | exo |
| Time to repopulation of a depopulated farm | Po(90) | exo |
| Maximum number of daily traced farms | Po(60) | exo |
| Probability of tracing an animal movement | Be(0.99) | exo |
| Probability of tracing a people movement | Be(0.40) | exo |
| Probability of tracing a vehicle transporting products movement | Be(0.70) | exo |
| Number of days for tracing | 60 | exo |

exo: CyL expert opinion, 2008.
$I(t)$: Number of infected animals in current farm at time $t$,
$N(t)$: Number of animals in current farm at time $t$.
Be$(p)$: Bernoulli distribution with parameter $p$; Po$(\lambda)$: Poisson distribution with parameter $\lambda$.

Table 3.3: Main parameters used in the Be-FAST algorithm and default values used for CSF case in Segovia [98, 99, 68].

## 3.3 Risk analysis

### 3.3.1 CSF epidemics in Bulgaria

#### Data and definitions

Bulgarian pig farms are categorized in five types based on (1) the level of biosecurity, (2) the trade patterns permitted and (3) the farm size [3]. The first type of farm is referred to as *Industrial*, which is characterized by high levels of biosecurity, no restrictions on pig trade and large number of pigs on farm. *Family farm type A* is the second type, similar to industrial farms in permitted trading but, usually, with a smaller farm size and lower level of biosecurity (i.e. medium instead of high). *Family farm type B* is the third type characterized by poor or no biosecurity, smaller farm size and pig trade only allowed to other non-industrial pig farms. The fourth type is the *Backyard* pig farm, which has poor or no biosecurity, a very small farm size (up to 5 pigs and no sows) and in which pig trade is not allowed (pigs are only for self-consumption). Finally, the last type of farm is the *East-Balkan* pig herd, which is managed traditionally (i.e. free-range pigs fed in open grass areas), has poor or no biosecurity level, usually has medium to small farm sizes and in which trade is only allowed to other East-Balkan pig herds.

Data used in this study consisted of detailed pig demographics and trade for each type of pig farm, which was provided by the Bulgarian Food Safety Agency. Specifically, the number of farms per municipality and per type of farm and the number of pigs per farm during 2010 were available (Figure 3.4). Pig movement records were also obtained and used to simulate CSF virus spread by direct contacts. Specifically, the farm of origin, the farm of destination, the day of shipment and the number of pigs shipped from January to October 2010 were used.

#### Modeling

The spread of CSF in Bulgaria both by direct and indirect contacts was modeled by adjusting the spatial and stochastic model Be-FAST for CSF. Note that local spread was defined here as the indirect CSF virus transmission by airborne spread or fomites from an infected farm to farms in close proximity ($< 2$ km) [98, 72]. In this study, model parameters were adapted to the Bulgarian conditions by using information from an expert opinion elicitation conducted on 31 May 2012 in Hannover and from the expert knowledge of Bulgarian Veterinary Authorities. Values of parameters that have changed from those described for the original model are detailed in Table 3.4.

A total of 5,000 different epidemics were run assuming different randomly selected index cases. Specifically, for the first 1,000 epidemics the index case was assumed to be a randomly selected backyard pig farm; for the next 1,000 epidemics the index case was assumed to be a family type B farm; for the next 1,000 epidemics, a family

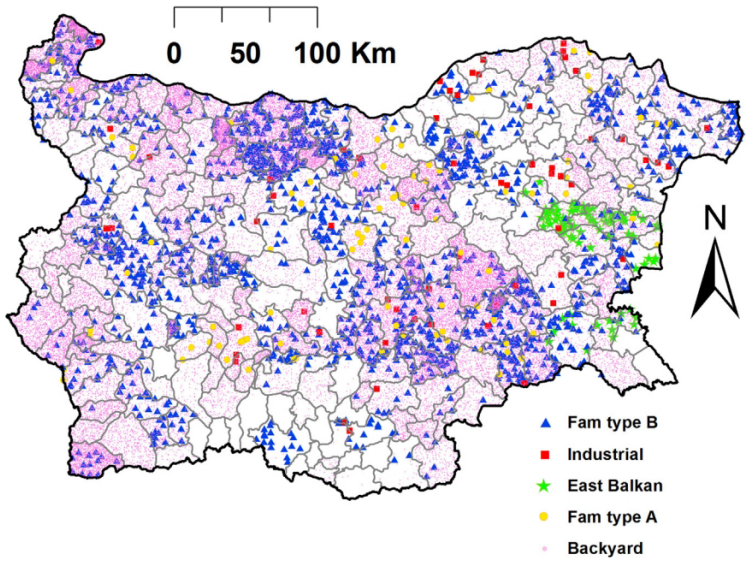


Figure 3.4: Spatial distribution of the different types of pig farms in Bulgaria during 2010.

| Parameter | Distribution or value per type of farm (I, A, B, BY, EA) | | | | |
|---|---|---|---|---|---|
| Within farm transmission ($\beta$) | 0.656 | 0.600 | 0.500 | 0.050 | 0.050 |
| Probability of infection due to infectious people | Be(0.0068) | Be(0.008) | Be(0.01) | Be(0.034) | Be(0.034) |
| Probability of infection due to infectious vehicles transporting products | Be(0.0065) | Be(0.0075) | Be(0.01) | Be(0.02) | Be(0.02) |
| Probability of restricting an animal movement outside the protection zones | Be(0.40) | Be(0.35) | Be(0.30) | Be(0.10) | Be(0.10) |
| Probability of tracing an animal movement | Be(0.99) | Be(0.95) | Be(0.90) | Be(0.30) | Be(0.30) |
| Probability of tracing a person movement | Be(0.67) | Be(0.60) | Be(0.50) | Be(0.10) | Be(0.10) |
| Probability of tracing a vehicle transporting products movement | Be(0.95) | Be(0.90) | Be(0.80) | Be(0.20) | Be(0.20) |

I: Industrial; A: family type A; B: family type B; BY: backyard; EB: East Balkan.
Be($p$): Bernoulli distribution with parameter $p$.

Table 3.4: Parameters different to the default ones implemented in the Be-FAST model used to simulate the CSF virus spread in Bulgaria.

type A farm; for the next 1,000 epidemics, an industrial farm and; for the last 1,000 epidemics an East Balkan pig herd.

Magnitude, duration and transmission patterns of the simulated CSF epidemics in Bulgaria were summarized using the median value, the mean value and the 95% probability intervals (PI). Values for the effective reproduction ratio of a farm $i$, $R(i)$, which was defined as the number of times that farm $i$ infects another farm in a *Susceptible* status considering all the simulations; and the risk of CSF virus introduction into a farm $i$, $Risk(i)$, which was defined as the number of times that farm $i$ becomes infected considering all the simulations were also calculated [7].

## Results

A total of 85% of the 5,000 simulations resulted in a non spread of the CSF (i.e. 85% of the intentionally infected index farms did not further spread the disease). Specifically, from those 1,000 simulations whose index farm was a backyard type, only 73 resulted in further spread of the disease. This number, when index cases were industrial, family type A, family type B and East Balkan pig herds were 387, 116, 86 and 105, respectively.

From those simulations that did result in CSF spread, the median and mean [95% PI] number of infected farms per epidemic was 1 and 2 [1; 4], respectively. The median and mean [95% PI] number of infected pigs were 2 and 371 [1; 4,477], respectively. The median and mean [95% PI] duration of the epidemic were 44 and 52 [17; 101] days, respectively. The median and mean [95% PI] number of infected Municipalities per epidemic were 1 and 1 [1; 3], respectively.

Simulations in which the index cases were industrial farms resulted in the largest epidemics (Figure 3.5A). Conversely, the magnitude and duration of the epidemics when index cases were backyards, family type A, family type B or East-Balkan pig herds were much smaller (Figure 3.5B).

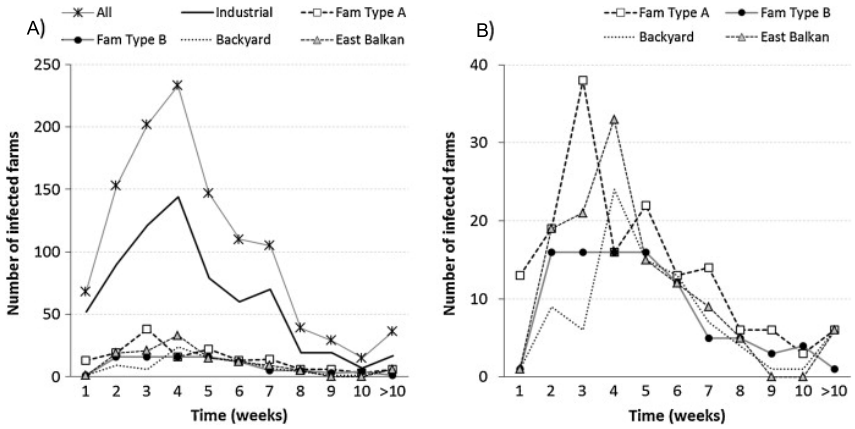


Figure 3.5: Epidemic curves of the 5,000 CSF epidemics in Bulgaria simulated with Be-FAST model. (a) Lines represent the number of infected farms per week when the index case is an industrial farm (first 1,000 simulations) or family type A (other 1,000 simulations) or family type B (other 1,000 simulations) or backyard (other 1,000 simulations) or East Balkan (other 1,000 simulations) or for all the 5,000 simulations. (b) Zoomed view for the epidemic curves of family type A, family type B, backyard and East Balkan.

The routes of infection in simulations in which the index cases were backyards were, primarily, vehicles (50.6%) and people (43.7%) and, in a very small proportion, local spread (5.7%). Contrarily, routes of infection in simulations in which the index cases were industrial farms were, mainly, local spread (61.8%) followed by animal movements (17.3%), vehicles (12.8%) and people (8.1%). Routes of infection for simulations in which the index cases were family type A farms were local spread

(28.9%), vehicles (28.8%), people (23.7%) and movement of animals (18.6%) and for simulations in which the index case were family type B farms were vehicles (49.5%), people (42.1%) and local spread (8.4%). Finally, routes of infection for simulations in which the index case were East-Balkan pig herds were vehicles (45.5%), people (41.3%) and local spread (13.2%). Detection of CSF outbreaks occurred, primarily (69%), by observation of clinical signs whereas zoning and tracing accounted for 21% and 10% of the detections, respectively.

A detailed description of the transmission routes and patterns after 5,000 simulations is presented in Tables 3.6 and 3.5. Detection of CSF outbreaks occurred, primarily (69%), by observation of clinical signs whereas zoning and tracing accounted for 21% and 10% of the detections, respectively. Most (95%) of the infected premises were backyards whereas most (56%) of the infectious premises were industrial farms. Industrial farms became infected very rarely (2.3%) and the source of infection was always another industrial farm. Family type B farms were also rarely infected (2.6%), but the source of infection was more diverse, including industrial, family type A and backyard farms. In the case of backyard farms, industrial farms were the main source of infection (52%) but all other types of farms also contributed to the CSF infection in backyards.

| Index case | Local spread | Animal movs. | Vehicle movs. | People movs. |
|---|---|---|---|---|
| Industrial (I) | 61.8% | 17.3% | 12.8% | 8.1% |
| Family type A (A) | 28.9% | 18.6% | 28.8% | 23.7% |
| Family type B (B) | 8.4% | 0.0% | 49.5% | 42.1% |
| Backyard (BY) | 5.7% | 0.0% | 50.6% | 43.7% |
| East Balkan (EA) | 13.2% | 0.0% | 45.5% | 41.3% |

Table 3.5: Main routes of CSF virus transmission from Transmission patterns indicated by using the percentage of infections regarding the 5,000 simulations of Be-FAST from (i.e. infectious) and to (i.e. infected) the different types of farms.Routes of transmission.

| Infectious farm type | I | A | B | BY | EB | Total |
|---|---|---|---|---|---|---|
| Industrial (I) | 2.3% | - | 2.1% | 51.7% | - | 51.6% |
| Family type A (A) | - | - | 0.3% | 12.1% | - | 12.4% |
| Family type B (B) | - | - | - | 7.9% | - | 7.9% |
| Backyard (BY) | - | - | 0.2% | 13.0% | - | 13.2% |
| East Balkan (EA) | - | - | - | 10.4% | - | 10.4% |
| Total | 2.3% | - | 2.6% | 94.1% | - | 100.0% |

Table 3.6: Transmission patterns indicated by using the percentage of infections regarding the 5,000 simulations of Be-FAST from (i.e. infectious) and to (i.e. infected) the different types of farms.

In general, the median [95% PI] of the $R$ and $Risk$ values for a pig farm in Bulgaria after 5,000 simulated CSF epidemics were 1 [1; 18] and 1 [1; 5.4]. As expected, industrial farms had the highest potential to infect other farms ($R = 7.5$) although they were concentrating a much lower risk of becoming infected ($Risk =$

2). The spatial distribution of the risk of CSF infection, which was very similar to the spatial distribution of $R$ (not shown), is presented in Figure 3.6.

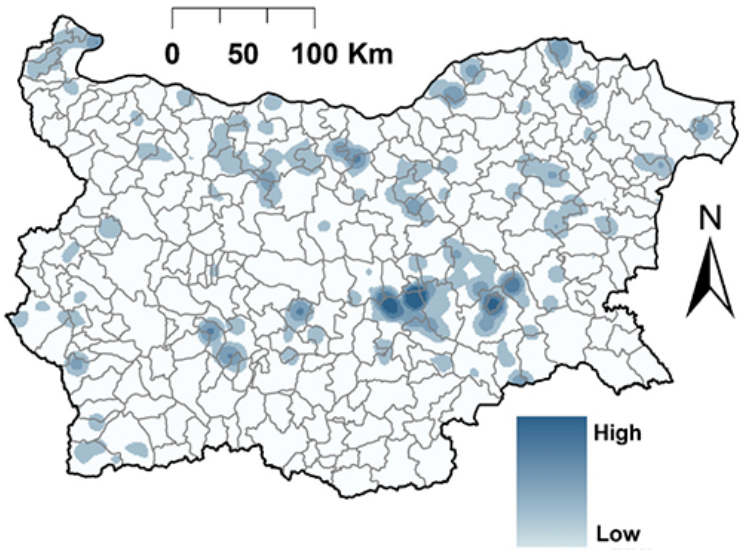


Figure 3.6: Spatial distribution of the risk of CSF infection in Bulgaria after 5,000 simulated epidemics. Map has been produced using the Kernel density function implemented in ArcGIS 9.3.

### 3.3.2 FMD epidemics in Peru

#### Data and definitions

No detailed updated animal census and spatial distribution of farms was available. We gathered information about the FMD susceptible livestock in Peru (i.e. cattle, pig, sheep, goat, alpaca and llama) at a district level (i.e. Peruvian lowest administrative unit) for year 1994 [63] and the evolution of different species available from 1994 to 2005 at a department level (union of districts) [66]. Nevertheless, animal demographics were believed to resemble that from the 1994-2005 census. We used linear regression models (fitted at $R^2 > 0.90$) to estimate the FMD susceptible species per district for the year 2013 using *lm* package in R language. The spatial distribution of farms was estimated by generating random points on a polygon using the *splancs* in R language. Farm size and species distribution were computed proportionally to year 1994 according to the estimation the FMD susceptible species in each district. Further information about input data estimations can be found in Martínez-López et al. (2014) [97]. The estimated total number of FMD susceptible animals and farms for Peru in 2013 were 27,671,887 and 1,986,750, respectively.

Information about animal movements within Peru from farm-to-farm was based on the Domestic Transit Health Certificate (CSTI) data provided by the National Agrarian Health Service of Peru (SENASA, Ministry of Agriculture). Specifically, we obtained information about the origin, destination, day of movement, number

of animals moved per species, purpose of the movement (i.e. trade, fattening, exhibition, reproduction, consumption, slaughterhouse, other), driver, owner and truck. It included 12,068 farms and 16,650 movements.

### Modeling

We adapted the Be-FAST model for the simulation of within- and between-farm FMD transmission in Peru, displayed in Table 3.7. We ran 1,000 epidemics using different index farms (i.e. different farm type, trade patterns on farm, farm size, and farm geolocation) because we were specifically interested in evaluating the magnitude and duration of potential FMD epidemics under diverse epidemiological conditions. We specifically defined four scenarios and run 250 epidemics for each one of them. In the first scenario, the index farm was randomly selected throughout Peru. In the second, index farms were selected from farms that had sent animals to other farms during 2011. For the third scenario, the index farm was randomly selected among those farms located in the two ecological regions of Peru, coast and mountains, which have the highest farm density. Finally, in the fourth scenario, the index farm was randomly selected only from farms located in northern Peru, which includes the districts currently not recognized as free-without vaccination (i.e. high risk region). Figure 3.7 shows the index farms selected for starting the FMD epidemics per scenario.

| Parameter | C | P | SG | AL | M |
|---|---|---|---|---|---|
| Within farm transmission ($\beta$) | 0.125 | 0.150 | 0.105 | 0.075 | 0.105 |
| Incubation period (days) | Po(2) | Po(2) | Po(2) | Po(2) | Po(2) |
| Probability of infection due to infectious vehicles transporting animals | Be(0.30) | Be(0.30) | Be(0.30) | Be(0.15) | Be(0.30) |
| Probability of infection due to infectious vehicles transporting products | Be(0.027) | Be(0.027) | Be(0.027) | Be(0.027) | Be(0.027) |
| Probability of infection due to infectious people | Be(0.021) | Be(0.021) | Be(0.021) | Be(0.021) | Be(0.021) |
| Probability of restricting an animal movement outside the protection zones | Be(0.40) | Be(0.35) | Be(0.30) | Be(0.10) | Be(0.10) |
| Duration of control zone (days) | 45 | 45 | 45 | 45 | 45 |
| Number of daily contacts with vehicles transporting products per farm | Po(0.1) | Po(0.1) | Po(0.1) | Po(0.1) | Po(0.1) |
| Number of daily contacts with people per farm | Po(0.3) | Po(0.3) | Po(0.3) | Po(0.3) | Po(0.3) |
| Daily probability of detection of the index case due to clinical signs | Be(0.047) | Be(0.047) | Be(0.047) | Be(0.047) | Be(0.047) |
| Daily probability of detection due to clinical signs | Be(0.0575) | Be(0.0575) | Be(0.0575) | Be(0.0575) | Be(0.0575) |
| Probability of restricting movements outside the protection zones | Be(0.30) | Be(0.30) | Be(0.30) | Be(0.10) | Be(0.30) |
| Duration of general movement restriction (days) | 90 | 90 | 90 | 90 | 90 |
| Maximum number of daily traced farms | 120 | 120 | 120 | 120 | 120 |
| Probability of tracing an animal movement | Be(0.80) | Be(0.80) | Be(0.80) | Be(0.40) | Be(0.80) |
| Probability of tracing a person movement | Be(0.70) | Be(0.70) | Be(0.70) | Be(0.35) | Be(0.70) |
| Probability of tracing a vehicle transporting products movement | Be(0.60) | Be(0.60) | Be(0.60) | Be(0.30) | Be(0.60) |

C: cattle; P: pig; SG: sheep and goat; AL: alpaca and llama; M: mixed.
Be($p$): Bernoulli distribution with parameter $p$; Po($\lambda$): Poisson distribution with parameter $\lambda$.

Table 3.7: Parameters different to the default ones implemented in the Be-FAST model used to simulate the FMD spread in Peru per type of animal premise (C, P, SG, AL, M).

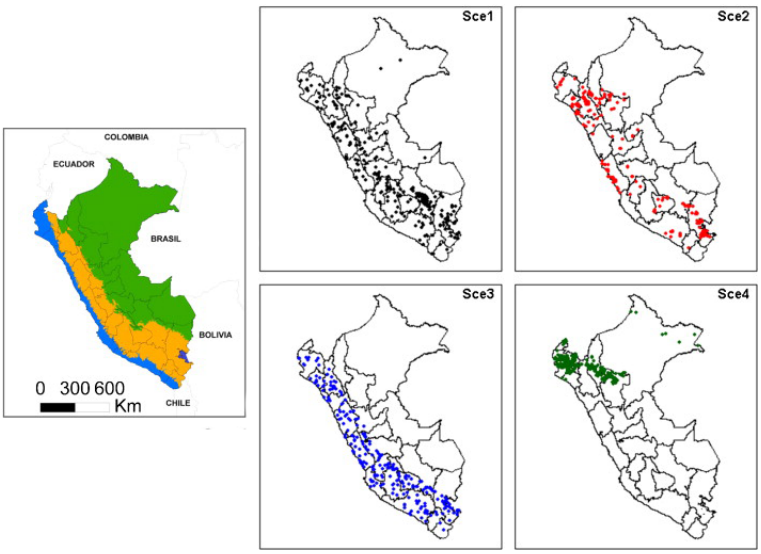


Figure 3.7: Ecological regions of Peru: coast (blue), mountains (orange), jungle (green) and location of the 1,000 index cases for the different scenarios: random selection of the index farm (sce1, n = 250), random index farm among those with outgoing shipments (sce2, n = 250), random index farm among those located in coast or mountain (sce3, n = 250) and random index farm among those located in the North of Peru (sce4, n = 250).

Model results were summarized using the mean and 95% probability interval (PI) and include: the number of infected and quarantined farms and animals; the duration (in days) of the epidemic and of the quarantine; the *Risk* of a farm becoming infected, defined as the number of times that a farm becomes infected considering all simulations; the effective reproduction ratio of a farm, $R$, defined as the amount of secondary infections caused by one infected farm in a *not naive* population [7]; the proportion of farms infected by direct contacts (i.e. animal movements), local spread, movements of people and of vehicles; the percentage of farms detected as infected by observation of clinical signs, active surveillance in the control and surveillance zones and tracing. We also computed the percentage of epidemics that resulted in a non-spread of the disease and that lasted more than 2 years (i.e. became endemic).

## Results

The overall mean [95% PI] number of infected farms and animals after 1,000 model simulations were 37 [1; 1,164] and 2,152 [0; 13,250], respectively. The mean [95% PI] time (in days) to detection and duration of the epidemic were 11 [1; 38] and 64 [0; 442], respectively. Spread of disease was primarily local with the mean [95% PI] number of infected districts equal to 4 [1; 16] and the mean [95% PI] distance from the source of infection to the infected farm of 4.5 km [0.5; 10.0]. Table 3.8 shows the outcomes summarized per scenario.

| | scenario 1 | scenario 2 | scenario 3 | scenario 4 | scenario 5 |
|---|---|---|---|---|---|
| Number of infected farms | 48 [0; 378] | 62 [0; 494] | 34 [0; 333] | 5 [0; 40] | 37 [0; 1,164] |
| Number of infected animals | 2,631 [0; 13,149] | 3,310 [9; 23,577] | 2,502 [0; 10,793] | 153 [0; 1,158] | 2,152 [0; 13,250] |
| Number of infected districts | 2 [1; 8] | 4 [1; 21] | 2 [1; 6] | 3 [1; 11] | 4 [1; 16] |
| Duration of the epidemic | 69 [0; 500] | 91 [0; 726] | 57 [0; 199] | 35 [0; 189] | 64 [0; 442] |
| Number of quarantined farms in the protection zone (<3 km) | 63,570 [332; 209,994] | 55,250 [390; 294,523] | 8,504 [8; 110,933] | 40,490 [11; 98,701] | 44,330 [135; 248,894] |
| Number of quarantined farms in the surveillance zone (<10 km) | 76,410 [52; 732,384] | 83,810 [47; 1,063,775] | 15,460 [114; 124,044] | 47,170 [153; 243,857] | 50,900 [11; 766,742] |
| Number of quarantined farms as suspected | 208 [1; 1853] | 225 [1; 2729] | 33 [1; 328] | 63 [1; 418] | 129 [1; 1,486] |
| Duration of quarantine | 115 [18; 717] | 120 [14; 730] | 98 [11; 364] | 77 [15; 219] | 103 [13; 718] |
| Time to detection | 9 [1; 31] | 10 [1; 37] | 10 [1; 24] | 15 [4; 49] | 11 [1; 38] |
| *Risk* | 1.09 [1; 2] | 1.83 [1; 9] | 1.10 [1; 2] | 1.11 [1; 2] | 1.32 [1; 4] |
| $R$ | 2.61 [1; 10] | 3.33 [1; 14] | 2.44 [1; 9] | 1.98 [1; 6] | 2.79 [1; 11] |

Table 3.8: Magnitude and duration of the FMD simulated epidemics within Peru when considering the overall 1,000 epidemics (total) and each of the scenarios (Mean [95% PI]).

The highest risk of a farm becoming FMD infected was concentrated in the coast and mountains ecological areas for all scenarios (Figure 3.8). Largest epidemics were observed when mixed farms were selected as index cases with a mean [95% PI] number of infected farms of 224 [1; 678]. This figure for index cases being cattle, sheep, goat, alpaca and llama farms was 192 [1; 1,993], 107 [1; 605], 48 [1, 589], 98 [1; 923], 73 [1; 289], 79 [3; 227], respectively. The mean [95% PI] *Risk* and $R$ values were 1.32 [1; 4] and 2.79 [1; 11], respectively.

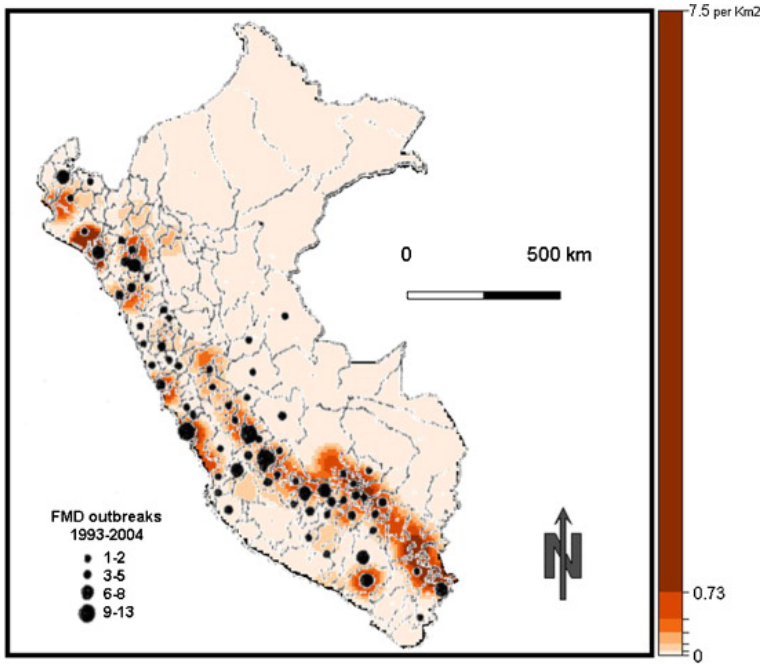


Figure 3.8: Comparison between the estimated overall risk map of the potential FMD spread in Peru (background color map) and the historical FMD outbreaks occurring in Peru from 1993 to 2004 (overlapped black dots [153]).

The percentage of farms (and percentage of animals) infected corresponding to small, medium and large farm sizes was 39.6% (8.4%), 54.5% (62.4%), and 5.9%

(29.2%), respectively. Up to 34.2% of the epidemics involved farms with only one animal species, which mainly were cattle (45%), pig (27%) and sheep (15%). Up to 12% of the epidemics involved farms with two species, 10% three, 14% four, 18% five, 7% six and 5% seven species. Farms infected in multi-species epidemics were mainly cattle (involved in 93% of the multi-species epidemics), sheep (90%), pig (80%), goat (59%) and mixed (44%) farms. Most (99.46%) of the infections were due to local spread, whereas 55.5% of detections were associated with active surveillance in the protection zone. Up to 8.5% of the initially infected index cases were not spreading further the disease. A total of 6% of the epidemics lasted more than 2 years (i.e. maximum simulated period). The highest magnitude and duration of FMD simulated epidemics were observed for scenario 2, whereas scenario 4 (random index farm among those located in northern Peru) had the lowest values. Table 3.9 shows the outcomes related to infection routes and detection summarized per scenario.

| | scenario 1 | scenario 2 | scenario 3 | scenario 4 | scenario 5 |
|---|---|---|---|---|---|
| % of infection by movement of animals | 0.01 | 0.21 | 0 | 0.02 | 0.09 |
| % of infection by local spread | 99.62 | 99.34 | 99.54 | 98.9 | 99.46 |
| % of infection by movement of people | 0.23 | 0.21 | 0.33 | 0.69 | 0.27 |
| % of infection by vehicles transporting products | 0.13 | 0.12 | 0.13 | 0.39 | 0.13 |
| % of infection by vehicles transporting animals | 0.01 | 0.12 | 0 | 0 | 0.05 |
| % of detection by clinical signs | 47.7 | 21.2 | 40.2 | 8.15 | 33.8 |
| % of detection by active surveillance in the control zone | 43.62 | 67.67 | 48.38 | 78.33 | 55.5 |
| % of detection by active surveillance in the surveillance zone | 4.08 | 4.11 | 4.96 | 13.31 | 4.57 |
| % of detection by tracing | 0.007 | 0.1 | 0 | 0 | 0.04 |
| % of non detected farms | 4.59 | 6.92 | 6.46 | 0.21 | 6.09 |
| % of epidemics lasting more than 2 years | 4.55 | 6.97 | 6.43 | 0.2 | 6.07 |
| % of epidemics not spreading further than the index case | 4.51 | 6.91 | 5.11 | 11.01 | 8.47 |

Table 3.9: Transmission routes and detection (in percentage) of the FMD simulated epidemics within Peru when considering the overall 1,000 epidemics (total) and each of the scenarios.

## 3.4 Economic epidemic estimation

### 3.4.1 Costs classification

The economic module operates at the end of each simulation day, Figure 3.3, in order to evaluate the daily cost of the given livestock disease epidemic in the specific area until eradication. Cost were classified according to the proposal by Saatkamp et al. [129]: (1) payable costs ($C_{\mathrm{p}}$), which are the costs paid directly by government authorities to control and eradicate the epidemic; (2) transferred costs ($C_{\mathrm{t}}$), which are the costs paid by the authorities to compensate other entities, such as farmers; (3) calculated costs ($C_{\mathrm{c}}$), which are losses generated in the livestock industry (e.g. transportation companies) that may occur between epidemic onset and eradication and that are not compensated by authorities; and (4) indirect costs suffered by the livestock industry as a result of the devaluation of meat.

### Payable costs

Human and material resources needed to apply the disease control measures, presented previously in Section 3.2, were considered payable costs since these resources should be paid directly by the authorities. This category included the following costs:

- $C_{\text{p,zn}}(i,t) \in \mathbb{R}$, which denotes the daily costs related to zoning around a detected farm $i$ on day $t$. Establishing protection areas where restricted activities are controlled requires administrative and security resources. It was computed as:

$$C_{\text{p,zn}}(i,t) = N_{\text{zn}}(i,t) \cdot MC_{\text{p,zn}}, \tag{3.4}$$

  where $N_{\text{zn}}(i,t) \in \mathbb{N}$ is the number of farms included in the zone around farm $i$ on day $t$ and not already included in another protection zone, and $MC_{\text{p,zn}} \in \mathbb{R}$ is the daily mean cost estimated to control one farm for one day.

- $C_{\text{p,cul}}(i,t) \in \mathbb{R}$ denotes the cost of culling and disinfecting farm $i$ on day $t$. It includes human and material resources as well as cleaning products needed during this process. It was computed as:

$$C_{\text{p,cul}}(i,t) = N_{\text{ani}}(i,t) \cdot MC_{\text{p,cul}}, \tag{3.5}$$

  where $N_{\text{ani}}(i,t) \in \mathbb{N}$ is the number of animals on farm $i$ on day $t$ and $MC_{\text{p,cul}} \in \mathbb{R}$ is the estimated mean cost per animal of culling and disinfecting the farm.

- $C_{\text{p,sm}}(i,t) \in \mathbb{R}$ denotes the cost of detecting possible infection on farm $i$ on day $t$. This cost involves sampling, laboratory analysis and employee salaries. In order to detect an infected farm $i$ on day $t$, random samples of size $N_{\text{sm}}(i,t) \in \mathbb{N}$ need to be collected and analysed. This samples were calculated using the formula extracted from Casal i Fàbrega and de Antonio [41]:

$$N_{\text{sm}}(i,t) = (1-(1-\alpha)^{1/P}) \cdot (N_{\text{ani}}(i,t) - \frac{P-1}{2}), \tag{3.6}$$

  where $P > 0$ is the expected prevalence of infected animals in the herd, and $\alpha$ is the required confidence level. Therefore, $C_{\text{p,sm}}(i,t)$ was computed as:

$$C_{\text{p,sm}}(i,t) = N_{\text{sm}}(i,t) \cdot MC_{\text{p,sm}}, \tag{3.7}$$

  where $MC_{\text{p,sm}} \in \mathbb{R}$ is the mean cost of testing one sample.

Taking into account previous costs, the total payable cost in one Monte-Carlo simulation is given by:

$$C_{\text{p}} = \sum_{t=1}^{T} \left( \sum_{i \in \Theta_{\text{zn}}(t)} C_{\text{p,zn}}(i,t) + \sum_{i \in \Theta_{\text{cul}}(t)} C_{\text{p,cul}}(i,t) + \sum_{i \in \Theta_{\text{sm}}(t)} C_{\text{p,sm}}(i,t) \right), \tag{3.8}$$

where $\Theta_{\text{cz}}(t)$ denotes the set of farms for which protection zones were set up on day $t$, $\Theta_{\text{cul}}(t)$, the set of farms culled on day $t$, and $\Theta_{\text{sm}}(t)$ the set of farms checked for infection on day $t$.

### Transferred costs

After culling animals, authorities usually compensate the affected livestock producers. These costs, so-called transferred costs, are denoted as $C_{\mathrm{t}} \in \mathbb{R}$, and they are strictly controlled by the authorities through a census of culled animals per outbreak. These costs are thought to cause significant economic impact [129]. Transferred costs was evaluated by considering:

$$C_{\mathrm{t}} = \sum_{t=1}^{T} \left( \sum_{i \in \Theta_{\mathrm{cul}}(t)} N_{\mathrm{ani}}(i,t) \cdot MC_{\mathrm{t,cul}}(i) \right), \tag{3.9}$$

where $MC_{\mathrm{t,cul}}(i) \in \mathbb{R}$ represents the compensation per animal depending on whether farm $i$ produced animals for fattening, farrowing or farrow-to-finish.

### Calculated costs

Losses borne by livestock companies during an epidemic are the most difficult to estimate [129]. In the present study, we considered the following costs:

- $C_{\mathrm{c,ds}}(i,t) \in \mathbb{R}$ denotes the daily cost of removing or destroying food and material on day $t$ on farm $i$ under quarantine. We assume that this cost is proportional to the number of animals per farm. It was computed as:

  $$C_{\mathrm{c,ds}}(i,t) = N_{\mathrm{ani}}(i,t) \cdot MC_{\mathrm{c,ds}}(i), \tag{3.10}$$

  where $MC_{\mathrm{c,ds}}(i) \in \mathbb{R}$ is the daily mean of material disposal per animal on farm $i$. This daily mean cost depends on whether farm $i$ produces animals for fattening, farrowing or farrow-to-finish.

- $C_{\mathrm{c,np}}(i,t) \in \mathbb{R}$ denotes the daily losses of farms due to animals culling. These losses occur during quarantine between detection and repopulation, when farms are not producing and so therefore are not deriving any benefit from the livestock. These losses were computed as:

  $$C_{\mathrm{c,np}}(i,t) = N_{\mathrm{ani}}(i,t) \cdot MC_{\mathrm{c,np}}, \tag{3.11}$$

  where $MC_{\mathrm{c,np}} \in \mathbb{R}$ is the mean daily cost due to non-production per day and per animal.

- $C_{\mathrm{c,tr}}(i,t) \in \mathbb{R}$, $C_{\mathrm{c,su}}(i,t) \in \mathbb{R}$ and $C_{\mathrm{c,vt}}(i,t) \in \mathbb{R}$ denote the daily losses caused by blockading farm $i$ in a protection zone and thereby eliminating the possibility of transporting livestock, animal supplies or veterinary services, respectively, that affect directly to transport companies or private veterinary services. These losses were computed as:

  $$C_{\mathrm{c,tr}}(i,t) = N_{\mathrm{tr}}(i,t) \cdot MC_{\mathrm{c,tr}}, \tag{3.12}$$

where $N_{\mathrm{tr}}(i,t) \in \mathbb{N}$ is the number of animals transported and blocked on farm $i$ on day $t$, and $MC_{\mathrm{c,tr}} \in \mathbb{R}$ is the mean cost of each movement.

$$C_{\mathrm{c,su}}(i,t) = N_{\mathrm{su}}(i,t) \cdot MC_{\mathrm{c,su}}, \tag{3.13}$$

where $N_{\mathrm{su}}(i,t) \in \mathbb{N}$ is the number of blocked movements of vehicles transporting supplies on farm $i$ on day $t$, and $MC_{\mathrm{c,su}} \in \mathbb{R}$ is the mean cost caused by blockading one supply movement.

$$C_{\mathrm{c,vt}}(i,t) = N_{\mathrm{vt}}(i,t) \cdot MC_{\mathrm{c,vt}}, \tag{3.14}$$

where $N_{\mathrm{vt}}(i,t) \in \mathbb{N}$ is the number of veterinarian services blocked on farm $i$ on day $t$, and $MC_{\mathrm{c,vt}} \in \mathbb{R}$ is the mean cost of blockading one veterinarian movement.

The total calculated cost of in one Monte-Carlo simulation is given by:

$$\begin{aligned} C_{\mathrm{c}} = \sum_{t=1}^{T} \Big( \sum_{i \in \Theta_{\mathrm{qt}}(t)} C_{\mathrm{c,ds}}(i,t) + \sum_{i \in \Theta_{\mathrm{zn}}(t)} \Big( C_{\mathrm{c,np}}(i,t) \\ + C_{\mathrm{c,tr}}(i,t) + C_{\mathrm{c,su}}(i,t) + C_{\mathrm{c,vt}}(i) \Big) \Big). \end{aligned} \tag{3.15}$$

where $\Theta_{\mathrm{qt}}(t)$ denotes the set of farms under quarantine on day $t$.

**Indirect costs**

Between detection of a putative CSF outbreak until its eradication, the meat price depreciates due to social alarm or trade restrictions, affecting the livestock market and its derivatives. In fact, all stakeholders fail to recover the expected benefits. In the present study, these so-called indirect losses, were denoted by $C_{\mathrm{i}}(t)$ and were computed daily as:

$$C_{\mathrm{i}}(t) = (MP_{\mathrm{obs}}(t) - MP_{\mathrm{pre}}(t)) \cdot MC_{\mathrm{i,tr}}(t) \cdot MC_{\mathrm{i,wg}}, \tag{3.16}$$

where $MC_{\mathrm{i,tr}}(t)$ is the number of animals traded at the end of day $t$, $MC_{\mathrm{i,wg}} \in \mathbb{R}$ mean animal weight, $MP_{\mathrm{obs}}(t)$ daily variation in the meat price observed during epidemic, and $MP_{\mathrm{pre}}(t) \in \mathbb{R}$ the predicted daily variation in meat price in the event that no epidemic occurs.

Thus, total indirect costs were computed as:

$$C_{\mathrm{i}} = \sum_{t=1}^{T} C_{\mathrm{i}}(t) \tag{3.17}$$

#### Total costs

Finally, Be-FAST was used to estimate the total costs as the sum of direct and indirect costs:

$$C_{\text{total}} = C_{\text{d}} + C_{\text{i}}, \tag{3.18}$$

where $C_{\text{d}} = C_{\text{p}} + C_{\text{t}} + C_{\text{c}}$ is called direct cost.

### 3.4.2 Numerical experiments

#### Model behaviour

We consider two numerical experiments based on data from the Spanish province of Segovia in 2008 and 2005. The objective of the first experiment is to evaluate the behaviour of the model in a recent swine sector framework. The second experiment aims to validate the model by comparing the outputs with past epidemics in Spain.

This first experiment was carried out with a database of Segovia in 2008, which consists in 1,400 farms, 1.11 million animals and 10,046 animal movements. The Be-FAST model was adapted to simulate the possible evolution of CSF epidemics and their economic impact for different epidemic magnitudes. For each case, we computed 1,000 Monte-Carlo simulations using a predetermined number of infected farms on the first day, $F = \{1, 5, 10, 15, 20, 25, 50, 75, 100\}$, expecting that such a numbers generate from mild to severe epidemics. For each series of epidemic magnitudes, we studied the behaviour of the economic losses for this disease (CSF) and region (Segovia) according to duration of the epidemic ($D$), the number of infected farms ($IF$) and the number of culled swine ($CS$). A first analysis explores the relationship between the costs (1) up to (4), defined in Section 3.4.1, through Spearman's correlation. Then, we estimate the best-fit regression equation ($R^2$, R-squared), among linear and quadratic, to estimate each cost depending on $D$, $IF$ and/or $CS$ variables.

#### Model validation

We point out that in 2006, an economic crisis in the swine industry in Spain changed the animal census and significantly reduced the farm density in Segovia [120]. We ran a second experiment for the specific case of the same disease (CSF) and region (Segovia) but with different farms distribution and movements behaviour; that is, data before 2006. Thus, we used a 2005 database with 2,354 farms, 1.41 million animals and 10,107 animal movements. Using this particular database, outcomes should be similar to the 1997-98 and 2000-01 epidemics data (see Table 3.1).

**Sensitivity analysis**

Finally, we enhanced the understanding of the relationships between input and output variables in Be-FAST, related to economic variables, through a sensitivity analysis. To do it, we repeated the first experiment with slight variations in the input parameters (±5%) identifying the ones that cause significant uncertainty in the main economic outputs.

### 3.4.3 Parameters estimation for CSF epidemics in Segovia

The parameters of the economic model, described in Section 3.4.1, were adapted to the specific case of CSF outbreaks in the Spanish province of Segovia. This choice of disease and region for validation allowed comparison with previous Be-FAST studies [99], though the present work was the first to have an economic focus. The values of the parameters (Tables 3.10 and 3.17) were estimated as follows:

- $MC_{\mathrm{p,zn}}$: We assumed that each farm in a protection zone needs to be tested daily by two workers (one for setting, controlling and checking the zoning area; and one for the administrative coordination). This parameter was estimated by averaging the daily gross salary of two employees.
- $MC_{\mathrm{p,cul}}$: We assumed that culling takes one day and we included salary for workers to cull and clean carcasses. We assumed that disinfection was performed twice [122]: once immediately after the culling process and again seven days later. This parameter is estimated from the daily gross salary of one worker and adding the average cost of chemical cleaning products described in the MAGRAMA report [122].
- $MC_{\mathrm{p,sm}}$: Expenses associated with sampling and laboratory analysis were obtained from the "Control program to eradicate the Aujeszky in Spain" [47], which followed a similar testing protocol (ELISA test).
- We assumed a disease prevalence of $P = 10\%$ and a confidence interval of $\alpha = 95\%$ [41].
- $MC_{\mathrm{t,cul}}(i)$: The estimated price per animal was taken from the official Spanish report [114], that provides an average cost per animal according to whether the farm produces animals for fattening, farrowing, or farrow-to-finish.
- $MC_{\mathrm{c,ds}}(i)$: The daily cost of feeding one swine was based on expert opinion of several companies specialised in swine feeding. In this approach, we assume seven day provisions per animal according to whether the farm produces animals for fattening, farrowing, or farrow-to-finish.
- $MC_{\mathrm{c,np}}$, $MC_{\mathrm{c,tr}}$, $MC_{\mathrm{c,su}}$ and $MC_{\mathrm{c,vt}}$: Costs associated with each blocked movement or service were based on information collected from active companies and professionals involved in the livestock sector.

- $MC_{\text{i,tr}}(t)$: This value was averaged from the traded swine database on day $t$; that is, the number of animals moved each day $t$ in a simulation.

- $MC_{\text{i,wg}}$: This value represents the average weight of one animal and was obtained from the Spanish government report [114].

| Parameter | $MC_{\text{p,zn}}$ | $MC_{\text{p,cul}}$ | $MC_{\text{p,sm}}$ | $MC_{\text{i,wg}}$ | $MC_{\text{t,cul}}(i)$ |
|---|---|---|---|---|---|
| Value(s) | 195.00 | 1.53 | 5.80 | 90.00 | {261.70, 320.25, 169.17}* |
| Units | €$^{\text{f,d}}$ | €$^{\text{a}}$ | €$^{\text{t}}$ | kg$^{\text{a}}$ | €$^{\text{a}}$ |
| Reference | exo | [122] | [114] | [114] | [47] |
| Parameter | $MC_{\text{c,np}}$ | $MC_{\text{c,tr}}$ | $MC_{\text{c,su}}$ | $MC_{\text{c,vt}}$ | $MC_{\text{c,ds}}(i)$ |
| Value(s) | 0.25 | 1.53 | 130.00 | 62.50 | {2.30, 2.30, 1.43}* |
| Units | €$^{\text{d,a}}$ | €$^{\text{a}}$ | €$^{\text{m}}$ | €$^{\text{m}}$ | €$^{\text{d,a}}$ |
| Reference | exo | exo | exo | exo | [122] |

*: fattening, farrowing, farrow-to-finish.
d: per day; f: per farm; a: per animal, m: per movement, t: per test.
exo: expert opinion from companies, professionals and experts.

Table 3.10: Parameter values of the economic module, presented in Section 3.4.1, which were used to guide experiments. Mean cost to control one farm for one day ($MC_{\text{p,zn}}$); mean cost per animal for culling and disinfecting a farm ($MC_{\text{p,cul}}$); mean cost for testing one sample ($MC_{\text{p,sm}}$); mean animal weight ($MC_{\text{i,wg}}$); mean cost for culling one animal of a farm $i$ according to whether the farm is of the type fattening, farrowing, or farrow-to-finish ($MC_{\text{t,cul}}(i)$); mean cost due to non-production per day and per animal ($MC_{\text{c,np}}$); mean cost of blockading one movement ($MC_{\text{c,tr}}$); mean cost of blockading one supply movement ($MC_{\text{c,su}}$); mean cost of blockading one veterinarian movement ($MC_{\text{c,vt}}$); mean cost of feeding one swine per day of a farm $i$ according to whether the farm is of the type fattening, farrowing, or farrow-to-finish ($MC_{\text{c,ds}}(i)$).

- $MP_{\text{obs}}(t)$ and $MP_{\text{pre}}(t) \in \mathbb{R}$: We analysed historical variation in the market price of swine (€/kg) during the 2001-02 CSF epidemic in Spain, which started in June 2001 and lasted 11 months [120], by drawing on data from two years before and two years after the epidemic in the Mercolleida database (`www.mercolleida.com`). Mercolleida is the Spanish government agency responsible for defining the price of swine each week. $MP_{\text{obs}}(t)$ denotes the observed meat prices during the studied period, while $MP_{\text{pre}}(t)$ denotes the predicted meat prices for the same period under the assumption of no epidemic. $MP_{\text{pre}}(t)$ was estimated using $MP_{\text{obs}}(t)$ before and after the CSF epidemic through the following backward operator:

$$\nabla MP_{\text{his}}(t) = \frac{MP_{\text{his}}(t)}{MP_{\text{his}}(t-1)}, \tag{3.19}$$

  where $MP_{\text{his}}(t)$ denotes the mean historical value of the meat price (€/kg) two years before and two years after the epidemic on the same day $t$ of the year excluding the year of the CSF epidemic. $MP_{\text{pre}}(t)$ was computed as:

$$MP_{\text{pre}}(t) = MP_{\text{pre}}(t-1) \cdot \nabla MP_{\text{his}}(t). \tag{3.20}$$

The initial value of the meat price at the beginning of the 2001-02 epidemic in Spain was 1.45 €/kg (`www.mercolleida.com`). Assuming the same initial value for $MP_{\text{pre}}(t)$ at the end of day $t = 1$, we found that daily values for $MP_{\text{obs}}(t)$ were usually below those of $MP_{\text{pre}}(t)$ (Figure 3.9). The lines for the two variables were parallel during the first and final weeks, whereas they showed funnel-shaped behaviour during the middle period of the epidemic. The results suggested several phases of cost behaviour:

- Initial period, before week 15: differences between $MP_{\text{obs}}(t)$ and $MP_{\text{pre}}(t)$ were around 0 while the standard deviation was $7 \cdot 10^{-3}$ (€/kg).
- Middle period, between week 15 and 35: differences between $MP_{\text{obs}}(t)$ and $MP_{\text{pre}}(t)$ increased progressively. The standard deviation reached $37 \cdot 10^{-3}$ (€/kg).
- Final period, after week 35: differences between $MP_{\text{obs}}(t)$ and $MP_{\text{pre}}(t)$ were around 0 while the standard deviation was $6.9 \cdot 10^{-3}$ (€/kg).

In this paper, we assumed a normal distribution (Kolmogorov-Smirnov normality test with significance value higher than 0.05) of $MP_{\text{obs}}(t)$ and $MP_{\text{pre}}(t)$ according to the three time-slots defined previously. Table 3.17 shows the mean and standard deviation estimated for $MP_{\text{obs}}(t)$ and $MP_{\text{pre}}(t)$ adapted to daily variation.

| index | phase | distribution |
|---|---|---|
| $MP_{\text{obs}}(t)$ | Initial period | Normal$(0.981, 3.14 \cdot 10^{-3})$ |
| $MP_{\text{obs}}(t)$ | Middle period | Normal$(1.003, 3.34 \cdot 10^{-3})$ |
| $MP_{\text{obs}}(t)$ | Final period | Normal$(1.013, 6.62 \cdot 10^{-3})$ |
| $MP_{\text{pre}}(t)$ | Initial period | Normal$(0.979, 4.10 \cdot 10^{-3})$ |
| $MP_{\text{pre}}(t)$ | Middle period | Normal$(0.987, 13.42 \cdot 10^{-3})$ |
| $MP_{\text{pre}}(t)$ | Final period | Normal$(1.012, 9.11 \cdot 10^{-3})$ |

Table 3.11: Estimated daily evolution of $MP_{\text{obs}}(t)$ and $MP_{\text{pre}}(t)$ in €/kg, according to the three time-slots defined in Section 3.4.3.

### 3.4.4 Results and discussion

#### Model behaviour

Table 3.12 summarises the results obtained for the considered values of $F$ and for each cost categories described in Section 3.4.1. The first cases (i.e. F={1,2}) involved a relative low number of infected farms ($IF$ < 10 farms) and culled swine ($CS$ <5,000 heads), and also a low duration of the epidemic ($D$ <60 days). On the contrary, last cases increased considerably these values (i.e. F={9,10} where

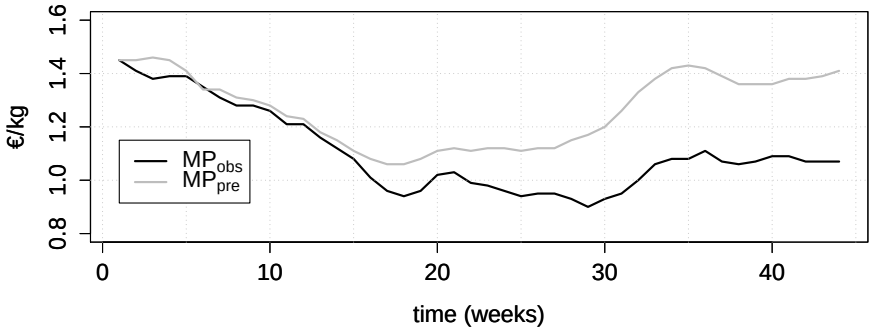


Figure 3.9: Variation in the observed swine price $MP_{\text{obs}}(t)$ (€/kg) and predicted swine price $MP_{\text{pre}}(t)$ (€/kg) during 2001-02 epidemic in Spain. Both indices were initially taken to be 1.45 €/kg, corresponding to $MP_{\text{obs}}(1)$.

$IF$ >100 farms, $CS$ >100,000 heads and $D$ >90 days). Similarly, economic variables described in Section 3.4.1 also increased with epidemic magnitude. The possible influence of $D$, $CS$ and $IF$ on economic variables showed a significant association with all types of economic costs (Table 3.14). In particular, indirect costs correlated strongly with epidemic duration ($\rho = 0.996$), payable costs correlated strongly with the number of infected farms ($\rho = 0.959$), and transferred and calculated costs were more correlated with the number of culled animals ($\rho = 0.996$ and 0.995, respectively) than with other variables. Furthermore, we have examined the total cost distribution among $C_{\text{p}}$, $C_{\text{t}}$ and $C_{\text{c}}$ as a function of the total direct cost (Table 3.12). From a general point of view, 77-83% of the direct costs were transferred costs, and calculated costs remained around 11% (Table 3.15). As direct costs increased, payable costs decreased from 11% to 4% from mild to severe epidemics.

Considering the Segovia-2008 scenario, the best-fit regression equations were:

$$C_{\text{i}}(D) = 0.998 - 0.031 \cdot D + 0.0003 \cdot D^2, \tag{3.21}$$

$$C_{\text{p}}(IF, CS) = 0.032 + 0.013 \cdot IF + 0.001 \cdot CS, \tag{3.22}$$

$$C_{\text{t}}(IF, CS, D) = -0.027 - 0.003 \cdot IF + 0.225 \cdot CS + 0.0005 \cdot D, \tag{3.23}$$

$$C_{\text{c}}(IF, CS, D) = -0.016 + 0.0005 \cdot IF + 0.033 \cdot CS + 0.0003 \cdot D. \tag{3.24}$$

These equations showed respective regression coefficients $R^2$ of 0.802, 0.988, 0.997 and 0.996, respectively. And the independent variables showed a significance value lower than 0.05; except for $D$ in Equation (3.23) where it reached 0.203 significance which implies that this variable may be excluded for the analysis (nevertheless, we have preferred to maintain it).

| Case | $F$ | $D$ | $CS$ | $IF$ | $C_{\text{total}}$ | $C_{\text{i}}$ | $C_{\text{d}}$ | $C_{\text{p}}$ | $C_{\text{t}}$ | $C_{\text{c}}$ |
|---|---|---|---|---|---|---|---|---|---|---|
| 1 | 1 | 54.5 | 1.9 | 2.42 | 0.91 | 0.36 | 0.55 | 0.06 | 0.42 | 0.06 |
| 2 | 5 | 76.9 | 8.3 | 10.49 | 3.18 | 0.84 | 2.34 | 0.22 | 1.84 | 0.27 |
| 3 | 10 | 82.8 | 15.9 | 19.58 | 5.28 | 0.86 | 4.43 | 0.38 | 3.52 | 0.52 |
| 4 | 15 | 87.6 | 23.6 | 28.72 | 7.41 | 0.87 | 6.53 | 0.53 | 5.23 | 0.77 |
| 5 | 20 | 89.4 | 30.3 | 36.67 | 9.33 | 0.98 | 8.35 | 0.65 | 6.71 | 0.98 |
| 6 | 25 | 90.3 | 36.6 | 44.59 | 11.09 | 1.00 | 10.09 | 0.76 | 8.13 | 1.19 |
| 7 | 50 | 91.5 | 69.4 | 82.34 | 20.00 | 1.19 | 18.80 | 1.19 | 15.36 | 2.25 |
| 8 | 75 | 90.7 | 100.7 | 119.18 | 28.25 | 1.19 | 27.06 | 1.43 | 22.36 | 3.27 |
| 9 | 100 | 94.7 | 132.1 | 156.56 | 36.55 | 1.38 | 35.16 | 1.56 | 29.32 | 4.28 |

Table 3.12: Results of economic modeling during the experiments described in Section 3.4.4. Values shown are averages for 1,000 simulations for each of the nine considered cases: Number of initial infected farms, $F$; duration in days, $D$; number of culled swine in miles, $CS$; number of infected farms, $IF$; total costs in millions (€), $C_{\text{total}}$; indirect costs in millions (€), $C_{\text{i}}$; direct costs in millions (€), $C_{\text{d}}$; payable costs indirect costs increased in millions (€), $C_{\text{p}}$; transferred costs in millions (€), $C_{\text{t}}$; and calculated costs in millions (€), $C_{\text{c}}$.

| index | phase | distribution |
|---|---|---|
| $MP_{\text{obs}}(t)$ | Initial period | Normal$(0.981, 3.14 \cdot 10^{-3})$ |
| $MP_{\text{obs}}(t)$ | Middle period | Normal$(1.003, 3.34 \cdot 10^{-3})$ |
| $MP_{\text{obs}}(t)$ | Final period | Normal$(1.013, 6.62 \cdot 10^{-3})$ |
| $MP_{\text{pre}}(t)$ | Initial period | Normal$(0.979, 4.10 \cdot 10^{-3})$ |
| $MP_{\text{pre}}(t)$ | Middle period | Normal$(0.987, 13.42 \cdot 10^{-3})$ |
| $MP_{\text{pre}}(t)$ | Final period | Normal$(1.012, 9.11 \cdot 10^{-3})$ |

Table 3.13: Estimated daily evolution of $MP_{\text{obs}}(t)$ and $MP_{\text{pre}}(t)$ in €/kg, according to the three time-slots defined in Section 3.4.3.

| Spearman's $\rho$ | $C_{\text{i}}$ | $C_{\text{p}}$ | $C_{\text{t}}$ | $C_{\text{c}}$ |
|---|---|---|---|---|
| $D$ | 0.629 | 0.718 | 0.647 | 0.651 |
| $IF$ | 0.462 | 0.959 | 0.901 | 0.900 |
| $CS$ | 0.419 | 0.896 | 0.996 | 0.995 |

Significance at the 0.01 level.

Table 3.14: Spearman's $\rho$ correlation analysis relating epidemic parameters and economic costs. Indirect costs, $C_{\text{i}}$; payable costs, $C_{\text{p}}$; transferred costs, $C_{\text{t}}$; calculated costs, $C_{\text{c}}$; duration of the epidemic, $D$; culled swine, $CS$; and number of infected farms, $IF$.

### Model validation

The model was validated using historical data from the 1997-98 CSF epidemic in Spain, which lasted 480 days and in which 99 outbreaks occurred and 609,147 animals were culled (Table 3.1). Total economic losses arising from the epidemic

| Case | $C_p(\%)$ | $C_t(\%)$ | $C_c(\%)$ |
|---|---|---|---|
| 1 | 11.61 | 77.10 | 11.29 |
| 2 | 9.44 | 78.97 | 11.58 |
| 3 | 8.62 | 79.64 | 11.74 |
| 4 | 8.19 | 80.03 | 11.78 |
| 5 | 7.87 | 80.34 | 11.79 |
| 6 | 7.62 | 80.54 | 11.85 |
| 7 | 6.33 | 81.68 | 11.98 |
| 8 | 5.28 | 82.63 | 12.09 |
| 9 | 4.45 | 83.38 | 12.17 |

Table 3.15: Relative contributions (in %) of payable costs, $C_p$, transferred costs, $C_t$, and calculated costs, $C_c$ to total direct cost. Percentages were calculated using the data in Table 3.12.

have been estimated at €89.5 million, of which €60 million (67%) were paid as transferred costs to compensate producers for culled animals [120]. Thus, €29.5 million (33%) were payable and calculated costs. Taking into account total costs and numbers of culled animals, we estimated that approximately €98.5 was paid in compensation per culled animal in the 1997-98 epidemic and €126.8 in the 2001-02 epidemic (Table 3.1). This indicates an increase in the price per animal.

Given the results in Table 3.12, dividing $C_t$ by $CS$, we obtained a mean compensation of €221 per culled animal in the Segovia epidemic simulations with 2008 database. Applying equations (3.22) and (3.24) to calculate payable and calculated costs, and assuming transferred costs to be €60 million, we estimated direct costs in that epidemic to be €81.52 million, which deviates by 9% from the actual value of €89.5 million. This deviation is quite reasonable given that numerous factors not included in our economic model may also influence costs, including the efficiency of control measures, livestock market performance in the region under analysis, farm density or the animal census.

The second experiment uses the 2005 database to obtain scenarios closer to available historical data, as explained in Section 3.4.2. We set $MC_{t,cul}(i)$ equal to €98.5 per animal, for every $i$, based on a cost of €60 million to 609,147 animals in the 1997-98 epidemic. Under this scenario, the best-fit regression equations were:

$$C_i(D) = -0.403 - 0.019 \cdot D + 0.0003 \cdot D^2, \tag{3.25}$$

$$C_p(IF, CS, D) = 0.049 + 0.001 \cdot IF + 0.005 \cdot CS + 0.0002 \cdot D, \tag{3.26}$$

$$C_t(CS) = 0.0985 \cdot CS, \tag{3.27}$$

$$C_c(IF, CS, D) = 0.028 - 0.001 \cdot IF + 0.035 \cdot CS + 0.0004 \cdot D. \tag{3.28}$$

These equations showed respective regression coefficients $R^2$ of 0.921, 0.804, 1.000 and 0.992, respectively. And the independent variables showed a significance value lower than 0.05.

On the basis of equations (3.26)-(3.28), direct costs of the epidemic were €85.26 million. Payable and calculated costs accounted for 30% of direct costs, while transferred costs accounted for 70%. These estimates are quite close to the corresponding observed values of 33% and 67% reported in del Pozo [120], as seen above.

The parameters used in this work (see Section 3.4.3) have been estimated with mean values in order to simplify the model but Be-FAST is easily adaptable to compute different estimators (as median, percentiles or even random distributions) according to type of farms, size of farms, locations or seasons.

As a third validation approach, we compared our modeling results based on data from the 2008 CSF epidemic in Spain with the results of economic modeling of CSF epidemics of different magnitudes simulated in Finland using Monte Carlo techniques [113]. In their analysis (Table 3.16), the authors distinguished between a normal epidemic, involving 1-5 infected farms, and a long epidemic, involving 6-33 infected farms. Our model estimated direct costs of approximately €0.55 million for a normal epidemic ($F = 1$) quite similar to the €0.5 million predicted in the Finnish study. In contrast, direct costs for severe epidemics ($F = 5, 10$) were estimated to be much higher in Spain €4.5 million than in Finland.

| case | normal epidemic | long epidemic |
|---|---|---|
| Lowest decrease in exporting | 1.4 | 5.4 |
| Medium decrease in exporting | 7.5 | 12.3 |
| Highest decrease in exporting | 13.2 | 19.2 |
| Direct costs | 0.5 | 1.6 |

Table 3.16: Predicted mean economic losses in millions (€) in Finland [113] as a result of a normal CSF epidemic (involving 1-5 infected farms) or a long epidemic (6-33 infected farms). Lowest, medium and highest decreases in exporting refer to a reduction of 10%, 50% and 100% in export demand, respectively.

The current cost classification is similar to the one introduced in Saatkamp et al. [129], which provides a losses dissociation depending on the nature of the affected entities within the swine industry. We note that similar economic studies developed a more accurate analysis regarding the meat demand in a country that derives in a better estimation of the indirect costs and the meat costumers benefits under price decrease in the meat market which may be influenced by international trade [113, 134, 116]. That kind of information was not available for our present work and should be included in a future analysis to enhance the economic module.

The outcomes presented in this work correspond to the specific case of CSF in Segovia (Spain) before and after 2006. Considering data from the 1997-98 epidemic in Spain, the validation experiments presented a reasonable approximation of the real economic losses. As we can see in the equations (3.21)-(3.28), the results change according to two different input databases (2008 and 2005) reflecting different pig sector frameworks. Thus, the Be-FAST model and the economic module could be transferable to diseases and regions in which conditions and parameters are similar to those studied in this work.

#### Sensitivity analysis

We examined the extent to which variations in our economic outputs affected the overall Be-FAST outputs. Previous work has already shown that Be-FAST is quite robust to variations in non-economic parameters [98]. Variations in economic parameters did not affect non-economic parameters of the epidemic; thus, changing economic parameters did not alter epidemic magnitude. Similarly, slight changes in most of the economic parameters did not significantly affect the final costs of the epidemic. However, only shown significant economic variation in $MC_{\mathrm{t,cul}}(i)$, which reducing or increasing by 5%, the direct costs varied up to 1.5% in mild epidemics and up to 1% in severe ones.

| index | phase | distribution |
|---|---|---|
| $MP_{\mathrm{obs}}(t)$ | Initial period | Normal$(0.981, 3.14 \cdot 10^{-3})$ |
| $MP_{\mathrm{obs}}(t)$ | Middle period | Normal$(1.003, 3.34 \cdot 10^{-3})$ |
| $MP_{\mathrm{obs}}(t)$ | Final period | Normal$(1.013, 6.62 \cdot 10^{-3})$ |
| $MP_{\mathrm{pre}}(t)$ | Initial period | Normal$(0.979, 4.10 \cdot 10^{-3})$ |
| $MP_{\mathrm{pre}}(t)$ | Middle period | Normal$(0.987, 13.42 \cdot 10^{-3})$ |
| $MP_{\mathrm{pre}}(t)$ | Final period | Normal$(1.012, 9.11 \cdot 10^{-3})$ |

Table 3.17: Estimated daily evolution of $MP_{\mathrm{obs}}(t)$ and $MP_{\mathrm{pre}}(t)$ in €/kg, according to the three time-slots defined in Section 3.4.3.

## 3.5 Conclusions

The two risk analysis carried out in Section 3.3 were one of the very first works that quantitatively describe the potential spread of CSF in backyard premises and FMD after eradication, respectively. As commented in the introduction, simulation tools such as Be-FAST allow to evaluate the evolution of diseases in areas where no previous epidemics have been reported or where the livestock framework has changed considerably since the last reported epidemic. Furthermore, it is important to note that studied territories (i.e. Bulgaria and Peru) have a small percentage of large farms (more than 1,000 animals) for industrial activity and a high density of small farms for local practices. This was an important motivation for the selection of a model able to simulate the spread within- and between farms. Moreover, based on the within-farm transmission process, Be-FAST allows to fit animal-to-animal effective contact rates according to the biosecurity level and the veterinarian practices associated to each type of holding. Thereby, the results allow to estimate the spread patterns and the magnitude of epidemics, as well as the effectiveness of the control measures. Additionally, the obtained results facilitate the identification of the spatial location of risk areas with high values of *Risk* and $R$ associated to each farm.

On the other hand, drawing on real economic data from the swine industry, we presented in Section 3.4 a spatio-temporal simulation model to estimate the

| +5% | $C_{\mathrm{p}}$ | $C_{\mathrm{t}}$ | $C_{\mathrm{c}}$ | $C_{\mathrm{d}}$ | $C_{\mathrm{i}}$ |
|---|---|---|---|---|---|
| $MC_{\mathrm{p,zn}}$ | 4.12; 4.92; 4.99 | - | - | 0.11; 0.69; 1.59 | - |
| $MC_{\mathrm{p,cul}}$ | 0.01; 0.05; 0.84 | - | - | 0.00; 0.01; 0.02 | - |
| $MC_{\mathrm{p,sm}}$ | 0.00; 0.00; 0.06 | - | - | 0.00; 0.00; 0.01 | - |
| $MC_{\mathrm{t,cul}}$ | - | 5.00; 5.00; 5.00 | - | 0.13; 1.63; 4.04 | - |
| $MC_{\mathrm{c,ds}}$ | - | - | 0.01; 0.22; 1.55 | 0.01; 0.10; 0.24 | - |
| $MC_{\mathrm{c,np}}$ | - | - | 0.00; 0.01; 0.10 | 0.00; 0.00; 0.03 | - |
| $MC_{\mathrm{c,tr}}$ | - | - | 0.14; 0.24; 0.33 | 0.03; 0.11; 0.23 | - |
| $MC_{\mathrm{c,su}}$ | - | - | 1.24; 1.73; 1.88 | 0.19; 0.79; 1.48 | - |
| $MC_{\mathrm{c,vt}}$ | - | - | 1.05; 1.49; 1.58 | 0.16; 0.67; 1.24 | - |
| $D$ | - | 0.01; 0.03; 1.99 | 0.01; 0.02; 1.26 | 0.01; 0.05; 1.67 | -0.05; 0.13; 0.17 |
| $IF$ | 1.43; 4.30; 4.56 | 0.01; 0.01; 0.07 | -0.19; 0.05; 0.21 | 0.15; 0.23; 0.59 | - |
| $CS$ | 0.03; 0.29; 0.42 | 5.01; 5.08; 6.14 | 4.25; 5.07; 7.45 | 3.35; 4.72; 4.81 | - |
| -5% | $C_{\mathrm{p}}$ | $C_{\mathrm{t}}$ | $C_{\mathrm{c}}$ | $C_{\mathrm{d}}$ | $C_{\mathrm{i}}$ |
| $D$ | - | - | - | - | -0.15; -0.11; 0.06 |

10th-Percentile; Median; 90th-Percentile.

Table 3.18: 10th-Percentile, Median and 90th-Percentile of the output variation percentage of the indirect costs ($C_{\mathrm{i}}$), payable costs ($C_{\mathrm{p}}$), transferred cost ($C_{\mathrm{t}}$) and calculated costs ($C_{\mathrm{c}}$) for the 9 cases presented in Table 3.12 with 1,000 simulations per case by changing ±5% the main epidemic variables and economic parameters: Mean cost to control one farm for one day ($MC_{\mathrm{p,zn}}$); mean cost per animal of culling and disinfecting a farm ($MC_{\mathrm{p,cul}}$); mean cost of testing one sample ($MC_{\mathrm{p,sm}}$); mean animal weight ($MC_{\mathrm{i,wg}}$); mean cost for culling one animal of a farm $i$ according to whether the farm is the type of fattening, farrowing, or farrow-to-finish ($MC_{\mathrm{t,cul}}(i)$); mean cost due to non-production per day and per animal ($MC_{\mathrm{c,np}}$); mean cost of each movement ($MC_{\mathrm{c,tr}}$); mean cost of blockading one supply movement ($MC_{\mathrm{c,su}}$); mean cost of blockading one veterinarian movement ($MC_{\mathrm{c,vt}}$); mean cost of feeding one swine per day of a farm $i$ according to whether the farm is the type of fattening, farrowing, or farrow-to-finish ($MC_{\mathrm{c,ds}}(i)$); duration of the epidemic (in days, $D$); number of infected farms ($IF$); number of culled swine (in miles, $CS$).

economic costs associated with CSF epidemics, which we validate using historical data on recent epidemics in Segovia, Spain. We compared how epidemic magnitude and duration affect predicted costs within the model, and compared these findings with actual data and predictions from another economic model. The strengths of our economic modeling approach are that we separately consider various types of costs, we assess the reliability of model parameters and we analyse how costs vary over the course of an epidemic. The key to model reliability is to include real economic data for the disease and region under study, which may be challenging to obtain because of the sensitive nature of the data. Well-informed models and data may improve the design strategies to control livestock diseases and minimise their global impact on animals and the economy. In future work, we plan to use our model to assess alternative control measures in order to minimise the global economic impact such as vaccination, preventive culling and risk-based surveillance.

# General conclusions and future work

This thesis documents the success of collaborative work between the mathematical and veterinary sciences. Joint research has allowed the development of competitive and useful tools to improve the knowledge on infectious diseases in very important areas of epidemiologist's duty, such as risk of introduction, early detection and risk of spread. These developments have required the ability to find efficient solutions to real needs and the ability to use the most appropriate methods within mathematics, statistics and computer science. The main achievements and potential future works are summarized below:

- The advection-deposition-survival model for Culicoides trajectory simulation, developed in Chapter 1 leads to the assessment of areas and periods at risk of introducing bluetongue virus in Spain from North Africa. The low computation cost required by this algorithm, the easy and fast numerical implementation and domain extrapolation makes this model a suitable tool for global vector-borne disease early warning system development.
- The automatic video-based disease detection system, developed in Chapter 2, has lead to the digital quantification of animals' motion. In addition, it has demonstrated that there is a significant reduction of movement in animals infected with African swine fever virus, and this reduction is detectable even before clinical signs. The high resolution of current video cameras and the use of computer vision algorithms aimed to object recognition and tracking will distinguish specific animal behaviors more precisely and thus improve the detection of febrile diseases.
- The model for disease epidemics simulation in the livestock sector, explained in Chapter 3, has allowed to study spread patterns and risk of classical swine fever and foot-and-mouth disease in Peru and Bulgaria. In addition, it showed additional economic losses estimation and classification for future epidemics of classical swine fever in Segovia (Spain). The adaptation to any notifiable diseases and new approaches for wildlife interaction will allow to work with new diseases such as African swine fever and more accurate spread outcomes in regions with high density of wild boar population, as in eastern and northeastern Europe.

# Annex: Publications

Veterinary Microbiology 165 (2013) 79–85

Contents lists available at SciVerse ScienceDirect

Veterinary Microbiology

journal homepage: www.elsevier.com/locate/vetmic

# Evaluation of the risk of classical swine fever (CSF) spread from backyard pigs to other domestic pigs by using the spatial stochastic disease spread model Be-FAST: The example of Bulgaria

Beatriz Martínez-López [a,b,*], Benjamin Ivorra [c], Angel Manuel Ramos [c], Eduardo Fernández-Carrión [c], Tsviatko Alexandrov [d], José Manuel Sánchez-Vizcaíno [a]

[a] VISAVET Center and Animal Health Department, Veterinary School, University Complutense of Madrid, Av. Puerta de Hierro s/n, 28040 Madrid, Spain
[b] IREC (CSIC-UCLM-JCCM), Ronda de Toledo s/n, 13005 Ciudad Real, Spain
[c] MOMAT Group, Applied Mathematics Department, University Complutense of Madrid, Spain
[d] Bulgarian Food Safety Agency, Blvd. Pencho Slaveikov 15A, Sofia, Bulgaria

ARTICLE INFO

*Article history:*
Received 30 October 2012
Received in revised form 17 January 2013
Accepted 30 January 2013

*Keywords:*
Spatial and stochastic simulation model
Classical swine fever
Backyard pigs
Bulgaria

ABSTRACT

The study presented here is one of the very first aimed at exploring the potential spread of classical swine fever (CSF) from backyard pigs to other domestic pigs. Specifically, we used a spatial stochastic spread model, called Be-FAST, to evaluate the potential spread of CSF virus (CSFV) in Bulgaria, which holds a large number of backyards (96% of the total number of pig farms) and is one of the very few countries for which backyard pigs and farm counts are available. The model revealed that, despite backyard pigs being very likely to become infected, infections from backyard pigs to other domestic pigs were rare. In general, the magnitude and duration of the CSF simulated epidemics were small, with a median [95% PI] number of infected farms per epidemic of 1 [1,4] and a median [95% PI] duration of the epidemic of 44 [17,101] days. CSFV transmission occurs primarily (81.16%) due to indirect contacts (i.e. vehicles, people and local spread) whereas detection of infected premises was mainly (69%) associated with the observation of clinical signs on farm rather than with implementation of tracing or zoning.

Methods and results of this study may support the implementation of risk-based strategies more cost-effectively to prevent, control and, ultimately, eradicate CSF from Bulgaria. The model may also be easily adapted to other countries in which the backyard system is predominant. It can also be used to simulate other similar diseases such as African swine fever.

© 2013 Elsevier B.V. All rights reserved.

## 1. Introduction

Classical swine fever (CSF) is a highly contagious viral disease of pigs and wild boars that causes severe economic impact due to the trade restrictions imposed on the affected countries. Considering that the pig sector in the European Union (EU) maintains a high level of production and exports, contributing more than €32 billion per year to

* Corresponding author at: VISAVET Center and Animal Health Department, Veterinary School, University Complutense of Madrid, Av. Puerta de Hierro s/n, 28040 Madrid, Spain. Tel.: +34 913943702; fax: +34 913943908.
*E-mail address:* beatriz@sanidadanimal.info (B. Martínez-López).

0378-1135/$ – see front matter © 2013 Elsevier B.V. All rights reserved.
http://dx.doi.org/10.1016/j.vetmic.2013.01.045

Preventive Veterinary Medicine 114 (2014) 47–63

Contents lists available at ScienceDirect

Preventive Veterinary Medicine

journal homepage: www.elsevier.com/locate/prevetmed

# A multi-analysis approach for space–time and economic evaluation of risks related with livestock diseases: The example of FMD in Peru

B. Martínez-López[a,b,c,*], B. Ivorra[d], E. Fernández-Carrión[d], A.M. Perez[a], A. Medel-Herrero[e], F. Sánchez-Vizcaíno[b], C. Gortázar[c], A.M. Ramos[d], J.M. Sánchez-Vizcaíno[b]

[a] Center for Animal Disease Modeling and Surveillance, VM: Medicine and Epidemiology, UC Davis, CA, USA
[b] VISAVET, Veterinary School, Complutense University of Madrid, Av. Puerta de Hierro s/n, 28040, Spain
[c] IREC (CSIC-UCLM-JCCM), Ronda de Toledo s/n, 13005 Ciudad Real, Spain
[d] Applied Mathematics Department, Mathematics School, Complutense University of Madrid, Plaza de Ciencias 3, 28040 Madrid, Spain
[e] Biomedical Research Network Center in Neurodegenerative Diseases, CIBERNED, ISCII, Madrid, Spain

ARTICLE INFO

*Keywords:*
Multi-analysis approach
Decision-support tool
Modeling
Social network analysis
Economic analysis
Foot-and-mouth disease
Peru

ABSTRACT

This study presents a multi-disciplinary decision-support tool, which integrates geo-statistics, social network analysis (SNA), spatial-stochastic spread model, economic analysis and mapping/visualization capabilities for the evaluation of the sanitary and socio-economic impact of livestock diseases under diverse epidemiologic scenarios. We illustrate the applicability of this tool using foot-and-mouth disease (FMD) in Peru as an example. The approach consisted on a flexible, multistep process that may be easily adapted based on data availability. The first module (mI) uses a geo-statistical approach for the estimation (if needed) of the distribution and abundance of susceptible population (in the example here, cattle, swine, sheep, goats, and camelids) at farm-level in the region or country of interest (Peru). The second module (mII) applies SNA for evaluating the farm-to-farm contact patterns and for exploring the structure and frequency of between-farm animal movements as a proxy for potential disease introduction or spread. The third module (mIII) integrates mI–II outputs into a spatial-stochastic model that simulates within- and between-farm FMD-transmission. The economic module (mIV) connects outputs from mI–III to provide an estimate of associated direct and indirect costs. A visualization module (mV) is also implemented to graph and map the outputs of module I–IV. After 1000 simulated epidemics, the mean (95% probability interval) number of outbreaks, infected animals, epidemic duration, and direct costs were 37 (1, 1164), 2152 (1, 13, 250), 63 days (0, 442), and US$ 1.2 million (1072, 9.5 million), respectively. Spread of disease was primarily local (<4.5 km), but geolocation and type of index farm strongly influenced the extent and spatial patterns of an epidemic. The approach is intended to support decisions in the last phase of the FMD eradication program in Peru, in particular to inform and support the implementation of risk-based surveillance and livestock insurance systems that may help to prevent and control potential FMD virus incursions into Peru.

© 2014 Elsevier B.V. All rights reserved.

* Corresponding author at: Center for Animal Disease Modeling and Surveillance, VM: Medicine and Epidemiology, UC Davis, CA, USA. Tel.: +1 530 752 7675.
*E-mail address:* beamartinezlopez@ucdavis.edu (B. Martínez-López).

0167-5877/$ – see front matter © 2014 Elsevier B.V. All rights reserved.
http://dx.doi.org/10.1016/j.prevetmed.2014.01.013

Preventive Veterinary Medicine 126 (2016) 66–73

Contents lists available at ScienceDirect

Preventive Veterinary Medicine

journal homepage: www.elsevier.com/locate/prevetmed

# Implementation and validation of an economic module in the Be-FAST model to predict costs generated by livestock disease epidemics: Application to classical swine fever epidemics in Spain

E. Fernández-Carrión[a,*], B. Ivorra[b], B. Martínez-López[c], A.M. Ramos[b], J.M. Sánchez-Vizcaíno[a]

[a] VISAVET Center and Animal Health Department, Veterinary School, Complutense University of Madrid, Av. Puerta de Hierro s/n, 28040 Madrid, Spain
[b] MOMAT Research Group, IMI-Institute and Applied Mathematics Department, Complutense University of Madrid, Plaza de Ciencias, 3, 28040 Madrid, Spain
[c] CADMS, Center for Animal Disease Modeling and Surveillance, School of Veterinary Medicine, UC Davis, One Shields Avenue, 1112 Tupper Hall, Davis, CA 95616, United States

ARTICLE INFO

*Article history:*
Received 21 May 2015
Received in revised form
11 December 2015
Accepted 14 January 2016

*Keywords:*
Epidemiological modeling
Economic modeling
Be-FAST
Classical swine fever
Risk surveillance

ABSTRACT

Be-FAST is a computer program based on a time-spatial stochastic spread mathematical model for studying the transmission of infectious livestock diseases within and between farms. The present work describes a new module integrated into Be-FAST to model the economic consequences of the spreading of classical swine fever (CSF) and other infectious livestock diseases within and between farms. CSF is financially one of the most damaging diseases in the swine industry worldwide. Specifically in Spain, the economic costs in the two last CSF epidemics (1997 and 2001) reached jointly more than 108 million euros. The present analysis suggests that severe CSF epidemics are associated with significant economic costs, approximately 80% of which are related to animal culling. Direct costs associated with control measures are strongly associated with the number of infected farms, while indirect costs are more strongly associated with epidemic duration. The economic model has been validated with economic information around the last outbreaks in Spain. These results suggest that our economic module may be useful for analysing and predicting economic consequences of livestock disease epidemics.

© 2016 Elsevier B.V. All rights reserved.

## 1. Introduction

Classical swine fever (CSF) is a highly contagious viral disease that affects wild and domestic swine, and it is considered one of the most economically damaging diseases in the swine industry worldwide (Boklund et al., 2009; Horst et al., 1999). Spain is the second largest producer of swine in the European Union (Eurostat, 2014). It was declared CSF-free in 1988, but two CSFV incursions have occurred since then: one in 1997–98 that affected the provinces of Lleida, Seville, Segovia and Saragossa; and another in 2001–02 that affected the provinces of Lleida, Castellón, Valencia, Cuenca and Barcelona. Both epidemics caused significant estimated economic costs (Table 1).

Although the characteristics and economic consequences of each CSF epidemic depend on the location, time period and type of holdings infected, strict control measures are always required in order to stop the spread of disease and eradicate it. Detection of a CSF epidemic in Europe triggers the following control measures, as mandated by Spanish and European Union regulations (MAGRAMA, 2011): (1) immediate culling of all swine on the infected farms and destruction of the carcasses; (2) restriction of movements related to the swine industry within CSF outbreak areas, including movements of animals, vehicles and people; (3) strict biosecurity measures, including disinfection of holdings, transport vehicles and other material at risk of contamination; (4) close monitoring and tracing to determine the infection source, with an emphasis on veterinary staff visits and movements of vehicles transporting animals and products; (5) zoning around infected holdings to accelerate the detection of infected farms in the vicinity and better control movements of vehicles that may spread disease.

All these control measures generate economic costs that are supported by government authorities and the swine industry. The

* Corresponding author.
*E-mail addresses:* eduardofc@ucm.es, eduardofernandezcarrion@gmail.com (E. Fernández-Carrión).

http://dx.doi.org/10.1016/j.prevetmed.2016.01.015
0167-5877/© 2016 Elsevier B.V. All rights reserved.

Transboundary and Emerging Diseases



ORIGINAL ARTICLE

# Early Detection of Infection in Pigs through an Online Monitoring System

M. Martínez-Avilés, E. Fernández-Carrión, J. M. López García-Baones and J. M. Sánchez-Vizcaíno

Visavet Centre and Animal Health Department, Faculty of Veterinary Sciences, Complutense University of Madrid, Madrid, Spain

**Keywords:**
surveillance; early detection; sentinel farms; sensor; motion; fever; African swine fever

**Correspondence:**
M. Martínez-Avilés. Visavet Centre and Animal Health Department, Faculty of Veterinary Sciences, Complutense University of Madrid, Avda. Puerta de Hierro, s/n, 28040, Madrid, Spain. Tel.: +34 91 394 37 02; Fax: +34 91 394 39 08;
E-mail: mmaviles@ucm.es

Received for publication October 31, 2014

doi:10.1111/tbed.12372

## Summary

Late detection of emergency diseases causes significant economic losses for pig producers and governments. As the first signs of animal infection are usually fever and reduced motion that lead to reduced consumption of water and feed, we developed a novel smart system to monitor body temperature and motion in real time, facilitating the early detection of infectious diseases. In this study, carried out within the framework of the European Union research project Rapidia Field, we tested the smart system on 10 pigs experimentally infected with two doses of an attenuated strain of African swine fever. Biosensors and an accelerometer embedded in an eartag captured data before and after infection, and video cameras were used to monitor the animals 24 h per day. The results showed that in 8 of 9 cases, the monitoring system detected infection onset as an increase in body temperature and decrease in movement before or simultaneously with fever detection based on rectal temperature measurement, observation of clinical signs, the decrease in water consumption or positive qPCR detection of virus. In addition, this decrease in movement was reliably detected using automatic analysis of video images therefore providing an inexpensive alternative to direct motion measurement. The system can be set up to alert staff when high fever, reduced motion or both are detected in one or more animals. This system may be useful for monitoring sentinel herds in real time, considerably reducing the financial and logistical costs of periodic sampling and increasing the chances of early detection of infection.

## Introduction

Early detection of emergency infectious diseases into a country free of those diseases is crucial to avoid the economic and health consequences of uncontrolled spread. Efforts to develop effective early detection systems in Europe increased after passive surveillance of several epizootics, including one involving classical swine fever (1997–98) in which delayed detection led to huge economic losses (Elbers et al., 1999; Stegeman et al., 2000). A widely used early detection approach is periodic sampling of vulnerable populations (sentinel sampling), in which a small number of herds are regularly sampled at the times and places considered at greatest risk of disease incursion (OIE, 2014).

While an effective approach, sampling is costly in economic and logistical terms and may not always be feasible. For example, the cost of wild bird sampling for avian influenza in 2006–2007 proved to be disproportionately high given the risk of infection in many European countries (Martinez et al., 2008). The need for sampling, and the costs involved, is even greater when emergency diseases such as African swine fever (ASF) are present or likely to be present. Sentinel surveillance can work well against ASF because the virus is detectable in blood as early as 30–48 h after infection and can remain detectable for weeks or months because neutralizing antibodies are not produced (Sánchez-Vizcaíno and Arias Neira, 2012). IgM or IgG can be detected from 4 to 8 days after a subacute infection has occurred and then last for several years (Sánchez-Vizcaíno and Arias Neira, 2012). However, the costs involved in sentinel testing can cause monitoring programmes to scale back because of budget restrictions

© 2015 Blackwell Verlag GmbH • Transboundary and Emerging Diseases.

# Annex: Further publications

PLOS ONE

RESEARCH ARTICLE

# Importance of Ecological Factors and Colony Handling for Optimizing Health Status of Apiaries in Mediterranean Ecosystems

**Irene Asensio[1]*, Marina Vicente-Rubiano[2,3], María Jesús Muñoz[1], Eduardo Fernández-Carrión[2,3], José Manuel Sánchez-Vizcaíno[2,3], Matilde Carballo[1]**

**1** Epidemiology & Environmental Health Department, Animal Health Research Center (CISA-INIA), Madrid, Spain, **2** VISAVET, Faculty of Veterinary Science, Universidad Complutense de Madrid (UCM), Madrid, Spain, **3** Animal Health Department, Faculty of Veterinary Science, Universidad Complutense de Madrid (UCM), Madrid, Spain

* irreasensio@gmail.com

CrossMark
click for updates

OPEN ACCESS

**Citation:** Asensio I, Vicente-Rubiano M, Muñoz MJ, Fernández-Carrión E, Sánchez-Vizcaíno JM, Carballo M (2016) Importance of Ecological Factors and Colony Handling for Optimizing Health Status of Apiaries in Mediterranean Ecosystems. PLoS ONE 11(10): e0164205. doi:10.1371/journal.pone.0164205

**Editor:** Bi-Song Yue, Sichuan University, CHINA

**Received:** April 4, 2016

**Accepted:** September 21, 2016

**Published:** October 11, 2016

**Copyright:** © 2016 Asensio et al. This is an open access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

**Data Availability Statement:** All relevant data are within the paper and its Supporting Information files.

**Funding:** This work was funded by RTA 2013-00042-C10-09 (JMSV, MJM) (Instituto Nacional de Investigación Agraria, INIA).

**Competing Interests:** The authors have declared that no competing interests exist.

## Abstract

We analyzed six apiaries in several natural environments with a Mediterranean ecosystem in Madrid, central Spain, in order to understand how landscape and management characteristics may influence apiary health and bee production in the long term. We focused on five criteria (habitat quality, landscape heterogeneity, climate, management and health), as well as 30 subcriteria, and we used the analytic hierarchy process (AHP) to rank them according to relevance. Habitat quality proved to have the highest relevance, followed by beehive management. Within habitat quality, the following subcriteria proved to be most relevant: orographic diversity, elevation range and important plant species located 1.5 km from the apiary. The most important subcriteria under beehive management were honey production, movement of the apiary to a location with a higher altitude and wax renewal. Temperature was the most important subcriterion under climate, while pathogen and *Varroa* loads were the most significant under health. Two of the six apiaries showed the best values in the AHP analysis and showed annual honey production of 70 and 28 kg/colony. This high productivity was due primarily to high elevation range and high orographic diversity, which favored high habitat quality. In addition, one of these apiaries showed the best value for beehive management, while the other showed the best value for health, reflected in the low pathogen load and low average number of viruses. These results highlight the importance of environmental factors and good sanitary practices to maximize apiary health and honey productivity.

## Introduction

Recent considerations about possible environmental factors contributing to the global decline in bee populations have implicated an array of causes, among which pests, pathogens, pesticides, nutrition and management have become the most important [1–4]. Poor nutrition

Acta Tropica 169 (2017) 163–169

Contents lists available at ScienceDirect

Acta Tropica

journal homepage: www.elsevier.com/locate/actatropica

# Risk mapping of West Nile virus circulation in Spain, 2015

Amaya Sánchez-Gómez[a,*], Carmen Amela[a], Eduardo Fernández-Carrión[b], Marta Martínez-Avilés[b], José Manuel Sánchez-Vizcaíno[b], María José Sierra-Moros[a]

[a] Coordinating Centre for Health Alerts and Emergencies, General Directorate of Public Health, Quality and Innovation, Ministry of Health, Social Services and Equality, Madrid, Spain Paseo del Prado 18-20, 28071 Madrid, Spain
[b] VISAVET Centre and Animal Health Department, Faculty of Veterinary Sciences, Complutense University, Avenida Puerta de Hierro, s/n, 28040 Madrid, Spain

ARTICLE INFO

*Article history:*
Received 17 November 2016
Received in revised form 10 February 2017
Accepted 13 February 2017
Available online 16 February 2017

*Keywords:*
West Nile virus
Risk assessment
Spain

ABSTRACT

West Nile fever is an emergent disease in Europe. The objective of this study was to conduct a predictive risk mapping of West Nile Virus (WNV) circulation in Spain based on historical data of WNV circulation. Areas of Spain with evidence of WNV circulation were mapped based on data from notifications to the surveillance systems and a literature review. A logistic regression-based spatial model was used to assess the probability of WNV circulation. Data were analyzed at municipality level. Mean temperatures of the period from June to October, presence of wetlands and presence of Special Protection Areas for birds were considered as potential predictors. Two predictors of WNV circulation were identified: higher temperature [adjusted odds ratio (AOR) 2.07, 95% CI 1.82–2.35, $p < 0.01$] and presence of wetlands (3.37, 95% CI 1.89–5.99, $p < 0.01$). Model validations indicated good predictions: area under the ROC curve was 0.895 (95% CI 0.870–0.919) for internal validation and 0.895 (95% CI 0.840–0.951) for external validation. This model could support improvements of WNV risk- based surveillance in Spain. The importance of a comprehensive surveillance for WNF, including human, animal and potential vectors is highlighted, which could additionally result in model refinements.

© 2017 Elsevier B.V. All rights reserved.

## 1. Introduction

West Nile virus (WNV) is a zoonotic flavivirus that is maintained in an enzootic cycle primarily through transmission between birds and ornithophilic (bird-biting) mosquitoes (Rossi et al., 2010). Mammals such as humans and horses are accidental dead-end hosts, as viremia resulting from WNV infection is insufficient to contribute to the amplification cycle (Jiménez-Clavero, 2012). Although up to 80% of human infections are asymptomatic, neuroinvasive disease occurs in less than 1% of the infections (Pérez-Ruiz et al., 2011).

West Nile fever (WNF) is an emergent disease in Europe, under surveillance in the European Union. Although serological surveys have pointed to WNV circulation in Europe since 1950s (Bardos et al., 1959), human outbreaks were relatively rare until recently. Since 2010, the WNV affected areas with confirmed human cases have further spread and the number of cases has increased markedly in Europe (ECDCD, 2011a, 2011b). Besides WNV lineage 1 that has traditionally circulated in Europe, lineage 2 has been recently detected in several European countries, including Greece (Papa et al., 2011) and Italy (Bagnarelli et al., 2011). In Spain, two cases of human WNF were confirmed in 2010 in Andalucía (southern Spain) (García-Bocanegra et al., 2011b). In the same area, outbreaks in horses have been detected since 2010–2014 (World Organisation for Animal Health, 2015).

Risk assessment (RA) is a systematic process for gathering, assessing and documenting information to assign a level of risk. This supports decision-making to reduce the negative consequences of public health risks (World Health Organization, 2012). Its importance has been particularly highlighted for assessing risks posed by emerging or re-emerging infections (Morgan et al., 2009). Determining spatial patterns is of particular interest to identify at risk territories. The objective of this study was to conduct a predictive risk mapping of WNV circulation in Spain based on historical data of WNV circulation, in order to support risk-based surveillance.

* Corresponding author.
*E-mail addresses:* amayasgomez@gmail.com (A. Sánchez-Gómez), camela@msssi.es (C. Amela), eduardofernandezcarrion@gmail.com (E. Fernández-Carrión), mmaviles@ucm.es (M. Martínez-Avilés), jmvizcaino@ucm.es (J.M. Sánchez-Vizcaíno), jsierra@msssi.es (M.J. Sierra-Moros).

http://dx.doi.org/10.1016/j.actatropica.2017.02.022
0001-706X/© 2017 Elsevier B.V. All rights reserved.

Received: 4 November 2016
DOI: 10.1111/tbed.12636

ORIGINAL ARTICLE

WILEY Transboundary and Emerging Diseases

# Understanding African Swine Fever infection dynamics in Sardinia using a spatially explicit transmission model in domestic pig farms

L. Mur[1] | J. M. Sánchez-Vizcaíno[2] | E. Fernández-Carrión[2] | C. Jurado[2] | S. Rolesu[3] | F. Feliziani[4] | A. Laddomada[3] | B. Martínez-López[5]

[1]Department of Diagnostic Medicine/Pathobiology, College of Veterinary Medicine, Kansas State University, Manhattan, KS, USA

[2]VISAVET Center and Animal Health Department, Veterinary School, Complutense University of Madrid, Madrid, Spain

[3]IZS della Sardegna, Centro di Sorveglianza Epidemiologica, Cagliari, Italy

[4]IZS delle Umbria e Marche, Perugia, Italy

[5]Center for Animal Disease Modeling and Surveillance (CADMS), University of California Davis, Davis, CA, USA

**Correspondence**
L. Mur, Department of Diagnostic Medicine/Pathobiology, College of Veterinary Medicine, Kansas State University, Manhattan, KS, USA.
Email: lmur@vet.k-state.edu

**Funding information**
This work was funded by the collaboration agreement "Ricerca, formazione e consulenza sulla Peste Suina Africana PSA" between the UCM and Regione Autonoma della Sardegna.

## Summary

African swine fever virus (ASFV) has been endemic in Sardinia since 1978, resulting in severe losses for local pig producers and creating important problems for the island's veterinary authorities. This study used a spatially explicit stochastic transmission model followed by two regression models to investigate the dynamics of ASFV spread amongst domestic pig farms, to identify geographic areas at highest risk and determine the role of different susceptible pig populations (registered domestic pigs, non-registered domestic pigs [*brado*] and wild boar) in ASF occurrence. We simulated transmission within and between farms using an adapted version of the previously described model known as Be-FAST. Results from the model revealed a generally low diffusion of ASF in Sardinia, with only 24% of the simulations resulting in disease spread, and for each simulated outbreak on average only four farms and 66 pigs were affected. Overall, local spread (indirect transmission between farms within a 2 km radius through fomites) was the most common route of transmission, being responsible for 98.6% of secondary cases. The risk of ASF occurrence for each domestic pig farm was estimated from the spread model results and integrated in two regression models together with available data for *brado* and wild boar populations. There was a significant association between the density of all three populations (domestic pigs, *brado*, and wild boar) and ASF occurrence in Sardinia. The most significant risk factors were the high densities of *brado* (OR = 2.2) and wild boar (OR = 2.1). The results of both analyses demonstrated that ASF epidemiology and infection dynamics in Sardinia create a complex and multifactorial disease situation, where all susceptible populations play an important role. To stop ASF transmission in Sardinia, three main factors (improving biosecurity on domestic pig farms, eliminating *brado* practices and better management of wild boars) need to be addressed.

KEYWORDS
African swine fever, between-farm-animal spatial transmission, disease spread model, endemic, risk factors, epidemiology

 © 2017 Blackwell Verlag GmbH 

# Bibliography


[1] *African Swine Fever*. Tech. rep. World Health Organitation for Animal Health, 2016.

[2] C. Alavani et al. "Modelling and simulation of a polluted water pumping process". In: *Mathematical and Computer Modelling* 51.5 (2010), pp. 461–472.

[3] T. Alexandrov, P. Kamenov, and K. Depner. "Surveillance and control of classical swine fever in Bulgaria, a country with a high proportion of non-professional pig holdings". In: *Epidémiol et santé anim* 59.60 (2011), pp. 140–142.

[4] R.M. Allen and H. Kanamori. "The potential for earthquake early warning in southern California". In: *Science* 300.5620 (2003), pp. 786–789.

[5] A. Allepuz et al. "Monitoring bluetongue disease (BTV-1) epidemic in southern Spain during 2007". In: *Preventive Veterinary Medicine* 96.3 (2010), pp. 263–271.

[6] T. Ancelle et al. "Outbreak of trichinellosis due to consumption of bear meat from Canada, France, September 2005". In: *Euro Surveill* 10.10 (2005), E051013.

[7] R. M Anderson, R.M. May, and B. Anderson. *Infectious diseases of humans: dynamics and control*. Vol. 28. Wiley Online Library, 1992.

[8] Dirección general de la ganadería. Subdirección general de sanidad animal. *Informe de la situación de la lengua azul en España*. Tech. rep. M.A.P.A. (Ministerio de Agricultura, Pesca y Alimentación), 2004 (7 de diciembre).

[9] Dirección general de la ganadería. Subdirección general de sanidad animal. *Informe de la situación de la lengua azul en España*. Tech. rep. M.A.P.A. (Ministerio de Agricultura, Pesca y Alimentación), 2005 (14 de enero).

[10] World Organisation for Animal Health. "Guide to terrestrial animal health surveillance". In: (2014).

[11] J.A. Backer and G. Nodelijk. "Transmission and control of African horse sickness in the Netherlands: a model analysis". In: *PLoS One* 6.8 (2011), e23066.

[12] T. Banhazi et al. "Practical problems associated with large scale deployment of PLF technologies on commercial farms". In: *Sessions of the 65th EAAP Annual Meeting*. 2014, pp. 105–112.

[13] C.T. Bauch. "The role of mathematical models in explaining recurrent outbreaks of infectious childhood diseases". In: *Mathematical Epidemiology*. Springer, 2008, pp. 297–319.

[14] K.J. Beven et al. "Testing a physically-based flood forecasting model (TOPMODEL) for three UK catchments". In: *Journal of Hydrology* 69.1-4 (1984), pp. 119–143.

[15] *Bluetongue*. Tech. rep. World Health Organitation for Animal Health, 2016.

[16] A. Boklund et al. "Comparing the epidemiological and economic effects of control strategies against classical swine fever in Denmark". In: *Preventive veterinary medicine* 90.3 (2009), pp. 180–193.

[17] F. Brauer. "Compartmental models in epidemiology". In: *Mathematical epidemiology*. Springer, 2008, pp. 19–79.

[18] F. Brauer et al. *Mathematical epidemiology*. Springer, 2008.

[19] Y. Braverman et al. "Dynamics of biting activity of C. imicola Kieffer (Diptera: Ceratopogonidae) during the year". In: *Israel Journal of Veterinary Medicine* 58.2/3 (2003), pp. 46–56.

[20] J.S. Brownstein, C.C. Freifeld, and L.C. Madoff. "Digital disease detection-harnessing the Web for public health surveillance". In: *New England Journal of Medicine* 360.21 (2009), pp. 2153–2157.

[21] M. Bruse and H. Fleer. "Simulating surface-plant-air interactions inside urban environments with a three dimensional numerical model". In: *Environmental Modelling & Software* 13.3 (1998), pp. 373–384.

[22] M. Bussiere and M. Fratzscher. "Towards a new early warning system of financial crises". In: *journal of International Money and Finance* 25.6 (2006), pp. 953–973.

[23] C. Calvete et al. "Modelling the distributions and spatial coincidence of bluetongue vectors Culicoides imicola and the Culicoides obsoletus group throughout the Iberian peninsula". In: *Medical and Veterinary Entomology* 22.2 (2008), pp. 124–134.

[24] I. Capua et al. "Newcastle disease outbreaks in Italy during 2000". In: *Veterinary Record* 150.18 (2002), pp. 565–568.

[25] Oficina Regional de la FAO para América Latina y el Caribe. *FAO reconoce a Perú esfuerzos para vencer la fiebre aftosa*. Tech. rep. Food and Agriculture Organization for the United Nations, 2013.

[26] H.A. Carneiro and E. Mylonakis. "Google trends: a web-based tool for real-time surveillance of disease outbreaks". In: *Clinical infectious diseases* 49.10 (2009), pp. 1557–1564.

[27] R.N. Charette. *Software engineering risk analysis and management*. Intertext Publications New York, 1989.

[28] A. Chedad et al. "AP—animal production technology: recognition system for pig cough based on probabilistic neural networks". In: *Journal of agricultural engineering research* 79.4 (2001), pp. 449–457.

[29] R. Clarke. *Building an Early Warning System in a Service Provider Network*. 2004.

[30] *Classical Swine Fever*. Tech. rep. World Health Organitation for Animal Health, 2016.

[31] N.J. Cook et al. "Infrared thermography detects febrile and behavioural responses to vaccination of weaned piglets". In: *animal* 9.02 (2015), pp. 339–346.

[32] S. Costard et al. "Epidemiology of African swine fever virus". In: *Virus research* 173.1 (2013), pp. 191–197.

[33] National Research Council. *Under the weather: climate, ecosystems, and infectious disease*. National Academies Press, 2001.

[34] A.C. Diego, P.J. Sánchez-Cordón, and J.M. Sánchez-Vizcaíno. "Bluetongue in Spain: from the first outbreak to 2012". In: *Transboundary and Emerging Diseases* 61.6 (2014), e1–e11.

[35] O. Diekmann and J.A.P. Heesterbeek. *Mathematical epidemiology of infectious diseases: model building, analysis and interpretation*. Vol. 5. John Wiley & Sons, 2000.

[36] M.A. Diuk-Wasser et al. "Modeling the spatial distribution of mosquito vectors for West Nile virus in Connecticut, USA". In: *Vector-Borne & Zoonotic Diseases* 6.3 (2006), pp. 283–295.

[37] D. Eagles et al. "Modelling spatio-temporal patterns of long-distance Culicoides dispersal into northern Australia". In: *Preventive veterinary medicine* 110.3 (2013), pp. 312–322.

[38] D.J.D. Earn. "A light introduction to modelling recurrent epidemics". In: *Mathematical epidemiology*. Springer, 2008, pp. 3–17.

[39] A.R.W. Elbers et al. "The classical swine fever epidemic 1997-1998 in the Netherlands: descriptive epidemiology". In: *Preventive veterinary medicine* 42.3 (1999), pp. 157–184.

[40] V. Exadaktylos et al. "Real-time recognition of sick pig cough sounds". In: *Computers and electronics in agriculture* 63.2 (2008), pp. 207–214.

[41] J. Casal i Fàbrega and E.M. de Antonio. *Problemas de epidemiología veterinaria*. Univ. Autònoma de Barcelona, 1999.

[42] Y. Fang et al. "Motion Based Animal Detection in Aerial Videos". In: *Procedia Computer Science* 92 (2016), pp. 13–17.

[43] H. Farid and E.P. Simoncelli. "Optimally rotation-equivariant directional derivative kernels". In: *International Conference on Computer Analysis of Images and Patterns*. Springer. 1997, pp. 207–214.

[44] H.Q. Feng et al. "Nocturnal windborne migration of ground beetles, particularly Pseudoophonus griseus (Coleoptera: Carabidae), in China". In: *Agricultural and Forest Entomology* 9.2 (2007), pp. 103–113.

[45] E. Fernández-Carrión et al. “Implementation and validation of an economic module in the Be-FAST model to predict costs generated by livestock disease epidemics: Application to classical swine fever epidemics in Spain”. In: *Preventive veterinary medicine* 126 (2016), pp. 66–73.

[46] *Foot-and-Mouth disease*. Tech. rep. World Health Organitation for Animal Health, 2016.

[47] Secretaría General de Agricultura y Alimentación. Dirección General de Ganadería. *Programa de control y erradicación de la enfermedad de Aujeszky en España*. Tech. rep. Ministerio de Medio Ambiente y Medio Rural y Marino, 2004.

[48] R. García-Lastra et al. “Bluetongue virus serotype 1 outbreak in the Basque Country (Northern Spain) 2007-2008. Data support a primary vector windborne transport”. In: *PLoS One* 7.3 (2012), e34421.

[49] M.G. Garner and S.A. Hamilton. “Principles of epidemiological modelling”. In: *Revue scientifique et technique (International Office of Epizootics)* 30.2 (2011), pp. 407–416.

[50] J.C. Gibbens et al. “Descriptive epidemiology of the 2001 foot-and-mouth disease epidemic in Great Britain: the first five months.” In: *The Veterinary Record* 149.24 (2001), pp. 729–743.

[51] J.J. Gibson. *The perception of the visual world*. Houghton Mifflin, 1950.

[52] J. Gloster et al. “Bluetongue in the United Kingdom and northern Europe in 2007 and key issues for 2008”. In: *Veterinary Record* 162.10 (2008), p. 298.

[53] J. Gloster et al. “Will bluetongue come on the wind to the United Kingdom in 2007?” In: *The Veterinary Record* 160.13 (2007), pp. 422–426.

[54] L. Gordis. *Epidemiology*. Elsevier Saunders, 1996.

[55] C. Guinat et al. “Paper: Effectiveness and practicality of control strategies for African swine fever: what do we really know?” In: *The Veterinary Record* 180.4 (2017), p. 97.

[56] J.N. Hanna et al. “An outbreak of Japanese encephalitis in the Torres Strait, Australia, 1995”. In: *Medical Journal of Australia* 165.5 (1996), pp. 256–261.

[57] C.K. Harris and P.L. Wiberg. “A two-dimensional, time-dependent model of suspended sediment transport and bed reworking for continental shelves”. In: *Computers & Geosciences* 27.6 (2001), pp. 675–690.

[58] G. Hendrickx et al. “A wind density model to quantify the airborne spread of Culicoides species during north-western Europe bluetongue epidemic, 2006”. In: *Preventive Veterinary Medicine* 87.1 (2008), pp. 162–181.

[59] M. Hilbert and P. López. “The world’s technological capacity to store, communicate, and compute information”. In: *science* 332.6025 (2011), pp. 60–65.

[60] B.K.P. Horn and B.G. Schunck. “Determining optical flow”. In: *Artificial intelligence* 17.1-3 (1981), pp. 185–203.

[61] H.S. Horst, R.B. Huirne, and A.A. Dijkhuizen. “Risks and economic consequences of introducing classical swine fever into The Netherlands by feeding swill to swine.” In: *Revue scientifique et technique (International Office of Epizootics)* 16.1 (1997), pp. 207–214.

[62] I. Iglesias et al. *Análisis probabilístico del riesgo potencial de entrada y difusión de la influenza aviar y de la enfermedad de Newcastle en España*. Tech. rep. Ministerio de Agricultura, Pesca y Alimentación, 2006.

[63] *III Censo Nacional Agropecuario 1994*. Tech. rep. Instituto Nacional de Estadística e Informática de Perú, 1994.

[64] *Impacts of animal disease outbreaks on livestock markets*. Tech. rep. Food and Agriculture Organization for the United Nations, 2006.

[65] H.J. In and S.U. Park. “A simulation of long-range transport of Yellow Sand observed in April 1998 in Korea”. In: *Atmospheric Environment* 36.26 (2002), pp. 4173–4187.

[66] Dirección General de Información agraria. *Capítulo 7 Estadística Pecuaria*. Tech. rep. Ministerio de Agricultura de Perú, 2013.

[67] *Informe de la situación de la lengua azul en España*. Tech. rep. M.A.P.A. (Ministerio de Agricultura, Pesca y Alimentación), 2004 (29 de octubre).

[68] B. Ivorra et al. “Mathematical formulation and validation of the Be-FAST model for Classical Swine Fever Virus spread between and within farms”. In: *Annals of operations research* 219.1 (2014), pp. 25–47.

[69] B. Ivorra et al. “Nonlinear Advection-Diffusion-Reaction Phenomena Involved in the Evolution and Pumping of Oil in Open Sea: Modeling, Numerical Simulation and Validation Considering the Prestige and Oleg Naydenov Oil Spill Cases”. In: *Journal of Scientific Computing* 70.3 (2015), pp. 1–27.

[70] A.W. Jalvingh et al. “Spatial and stochastic simulation to evaluate the impact of events and control measures on the 1997-1998 classical swine fever epidemic in The Netherlands. I. Description of simulation model”. In: *Preventive veterinary medicine* 42.3 (1999), pp. 271–295.

[71] S.M. Karlsson and J. Bigun. “Lip-motion events analysis and lip segmentation using optical flow”. In: *2012 IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops*. IEEE. 2012, pp. 138–145.

[72] S. Karsten, G. Rave, and J. Krieter. “Monte Carlo simulation of classical swine fever epidemics and control: I. General concepts and description of the model”. In: *Veterinary Microbiology* 108.3 (2005), pp. 187–198.

[73] B.H. Kay, I.D. Fanning, and P. Mottram. “The vector competence of Culex annulirostris, Aedes sagax and Aedes alboannulatus for Murray Valley encephalitis virus at different temperatures”. In: *Medical and veterinary entomology* 3.2 (1989), pp. 107–112.

[74] M.J. Keeling and K.T.D. Eames. “Networks and epidemic models”. In: *Journal of the Royal Society Interface* 2.4 (2005), pp. 295–307.

[75] W.O. Kermack and A.G. McKendrick. “A contribution to the mathematical theory of epidemics”. In: *Proceedings of the Royal Society of London A: mathematical, physical and engineering sciences*. Vol. 115. 772. The Royal Society. 1927, pp. 700–721.

[76] A.S. Khobragade, K.D. Kulat, and C.G. Dethe. “Motion analysis in video using optical flow techniques”. In: *International Journal of Information Technology and Knowledge Management* 5.1 (2012), pp. 9–12.

[77] S. Khomenko et al. “African swine fever in the Russian Federation: risk factors for Europe and beyond”. In: *Empres Watch* 28 (2013), pp. 1–14.

[78] D.P. King et al. “Development of a TaqMan® PCR assay with internal amplification control for the detection of African swine fever virus”. In: *Journal of virological methods* 107.1 (2003), pp. 53–61.

[79] I. Kleinschmidt et al. “Use of generalized linear mixed models in the spatial analysis of small-area malaria incidence rates in KwaZulu Natal, South Africa”. In: *American Journal of Epidemiology* 153.12 (2001), pp. 1213–1221.

[80] D. Klinkenberg et al. “Within-and between-pen transmission of classical swine fever virus: a new method to estimate the basic reproduction ratio from transmission experiments”. In: *Epidemiology and infection* 128.02 (2002), pp. 293–299.

[81] F. Koenen et al. “Epidemiological characteristics of an outbreak of classical swine fever in an area of high pig density”. In: *Veterinary Record* 139.15 (1996), pp. 367–371.

[82] S. Krauß. “Microscopic modeling of traffic flow: Investigation of collision free vehicle dynamics”. In: (1998).

[83] V.V. Krzhizhanovskaya et al. “Flood early warning system: design, implementation and computational modules”. In: *Procedia Computer Science* 4 (2011), pp. 106–115.

[84] J. Kuhlemann et al. “Regional synthesis of Mediterranean atmospheric circulation during the last glacial maximum”. In: *Science* 321.5894 (2008), pp. 1338–1340.

[85] F. Le Gall. *Economic and social consequences of animal diseases*. 2009.

[86] J.D. Lee et al. “Collision warning timing, driver distraction, and driver response to imminent rear-end collisions in a high-fidelity driving simulator”. In: *Human factors* 44.2 (2002), pp. 314–334.

[87] K. Lee, A. Agrawal, and A. Choudhary. “Real-time disease surveillance using twitter data: demonstration on flu and cancer”. In: *Proceedings of the 19th ACM SIGKDD international conference on Knowledge discovery and data mining*. ACM. 2013, pp. 1474–1477.

[88] D.J. Leprince et al. “Body size of Culicoides variipennis (Diptera: Ceratopogonidae) in relation to bloodmeal size estimates and the ingestion of Onchocerca cervicalis (Nematoda: Filarioidea) microfiliariae”. In: *Journal of the American Mosquito Control Association* (1989).

[89] T.H. Lillie, W.C. Marquardt, and R.H. Jones. "The flight range of Culicoides variipennis (Diptera: Ceratopogonidae)". In: *The Canadian Entomologist* 113.05 (1981), pp. 419–426.

[90] N.M. Lind et al. "Validation of a digital video tracking system for recording pig locomotor behaviour". In: *Journal of neuroscience methods* 143.2 (2005), pp. 123–132.

[91] S. Lloyd. "Least squares quantization in PCM". In: *IEEE transactions on information theory* 28.2 (1982), pp. 129–137.

[92] D. Lupulovic et al. "First serological study of hepatitis E virus infection in backyard pigs from Serbia". In: *Food and Environmental Virology* 2.2 (2010), pp. 110–113.

[93] S. Ma and Y. Xia. *Mathematical understanding of infectious disease dynamics*. Vol. 16. World Scientific, 2009.

[94] S. Maan et al. "Complete genome characterisation of a novel 26th bluetongue virus serotype from Kuwait". In: *PLoS One* 6.10 (2011), e26147.

[95] L.P. Macfadyen and S. Dawson. "Mining LMS data to develop an "early warning system" for educators: A proof of concept". In: *Computers & education* 54.2 (2010), pp. 588–599.

[96] T.N. Madsen and A.R. Kristensen. "A model for monitoring the condition of young pigs by their drinking behaviour". In: *Computers and electronics in agriculture* 48.2 (2005), pp. 138–154.

[97] B. Martínez-López et al. "A multi-analysis approach for space-time and economic evaluation of risks related with livestock diseases: The example of FMD in Peru". In: *Preventive veterinary medicine* 114.1 (2014), pp. 47–63.

[98] B. Martínez-López et al. "A novel spatial and stochastic model to evaluate the within-and between-farm transmission of classical swine fever virus. I. General concepts and description of the model". In: *Veterinary microbiology* 147.3 (2011), pp. 300–309.

[99] B. Martínez-López et al. "A novel spatial and stochastic model to evaluate the within and between farm transmission of classical swine fever virus: II Validation of the model". In: *Veterinary microbiology* 155.1 (2012), pp. 21–32.

[100] B. Martínez-López et al. "BTV: Risk of introduction into Spain associated with wind streams". In: *12th Symposium of the International Society for Veterinary Epidemiology and Economics*. 2009, pp. 94–96.

[101] M. Martínez-Avilés et al. "Early detection of infection in pigs through an online monitoring system". In: *Transboundary and emerging diseases* (2015).

[102] E. Massad et al. *Fuzzy logic in action: Applications in epidemiology and beyond*. Vol. 232. Springer Science & Business Media, 2009.

[103] S.G. Matthews et al. "Early detection of health and welfare compromises through automated detection of behavioural changes in pigs". In: *The Veterinary Journal* 217 (2016), pp. 43–51.

[104] M. McLaws et al. “Factors associated with the early detection of foot-and-mouth disease during the 2001 epidemic in the United Kingdom”. In: *The Canadian Veterinary Journal* 50.1 (2009), p. 53.

[105] C.A. Mebus. “African swine fever”. In: *Advances in virus research* 35 (1988), pp. 251–269.

[106] M. Melero et al. “Thermal reference points as an index for monitoring body temperature in marine mammals”. In: *BMC research notes* 8.1 (2015), p. 411.

[107] P.S. Mellor, J. Boorman, and M. Baylis. “Culicoides biting midges: their role as arbovirus vectors”. In: *Annual review of entomology* 45.1 (2000), pp. 307–340.

[108] G. Milne, C. Fermanis, and P. Johnston. “A mobility model for classical swine fever in feral pig populations”. In: *Veterinary research* 39.6 (2008), p. 1.

[109] V. Moennig. “Introduction to classical swine fever: virus, disease and control policy”. In: *Veterinary microbiology* 73.2 (2000), pp. 93–102.

[110] S.S. Morse. *Emerging viruses*. Oxford University Press on Demand, 1996.

[111] J. Naranjo and O. Cosivi. “Elimination of foot-and-mouth disease in South America: lessons and challenges”. In: *Philosophical Transactions of the Royal Society of London B: Biological Sciences* 368.1623 (2013), p. 20120381.

[112] S. Nickovic et al. “A model for prediction of desert dust cycle in the atmosphere”. In: *Journal of geophysical research* 106.D16 (2001), pp. 18113–18129.

[113] J.K. Niemi et al. “Simulated financial losses of classical swine fever epidemics in the Finnish pig production sector”. In: *Preventive veterinary medicine* 84.3 (2008), pp. 194–212.

[114] *Orden ARM/293/2011, de 7 de febrero*. Tech. rep. Boletín Oficial del Estado, 2011.

[115] *Pastos naturales*. Tech. rep. Ministerio de Agricultura y Riego de Perú, 2013.

[116] D.L. Pendel et al. “Economic assessment of FMDv releases from the National Bio and Agro Defense Facility”. In: *Plos One* 10 (2015). Ed. by Public Library of Science.

[117] M.L. Penrith and D. Nyakahuma. *Recognizing African swine fever: a field manual*. Tech. rep. 9. Food & Agriculture Org., 2000.

[118] M. Porta. *A dictionary of epidemiology*. Oxford university press, 2014.

[119] E. Pozio et al. “Development of harmonised schemes for the monitoring and reporting of Trichinella in animals and foodstuffs in the European Union”. In: (2010).

[120] M. del Pozo Vegas. “Estudio del brote de Peste Porcina Clásica ocurrido en Castilla y León en los años 1997/1998”. In: (2006).

[121] Dirección General de Sanidad de la Producción Agraria. *Manual práctico de actuaciones contra la PPC*. Tech. rep. Ministerio de Agricultura, Alimentación y Medio Ambiente, 2006.

[122] Dirección General de Sanidad de la Producción Agraria. *Manual práctico de actuaciones contra la PPC*. Tech. rep. Ministerio de Agricultura, Alimentación y Medio Ambiente, 2011.

[123] *Report on the outcome of the EU co-financed animal disease eradication and monitoring programmes in the MS and the EU as a whole*. Tech. rep. Food Chain Evaluation Consortium, 2011.

[124] *Rift Valley Fever*. Tech. rep. World Health Organitation for Animal Health, 2016.

[125] J.L. Ritchie-Dunham and J.F. Méndez Galván. "Evaluating epidemic intervention policies with systems thinking: a case study of dengue fever in Mexico". In: *System Dynamics Review* 15.2 (1999), pp. 119–138.

[126] D.J. Rogers and M.J. Packer. "Vector-borne diseases, models, and global change". In: *The Lancet* 342.8882 (1993), pp. 1282–1284.

[127] J.R. Romero, L.C. Villamil, and J.A. Pinto. "Impacto económico de enfermedades animales en sistemas productivos en Sudamerica: estudios de caso". In: *Revue scientifique et technique-Office international des épizooties* 18.2 (1999), pp. 498–511.

[128] M.O. Ruiz et al. "Local impact of temperature and precipitation on West Nile virus infection in Culex species mosquitoes in northeast Illinois, USA". In: *Parasites & vectors* 3.1 (2010), p. 19.

[129] H.W. Saatkamp, P.B.M. Berentsen, and H.S. Horst. "Economic aspects of the control of classical swine fever outbreaks in the European Union". In: *Veterinary microbiology* 73.2 (2000), pp. 221–237.

[130] H.W. Saatkamp et al. "Simulation studies on the epidemiological impact of national identification and recording systems on the control of classical swine fever in Belgium". In: *Preventive Veterinary Medicine* 26.2 (1996), pp. 119–132.

[131] J.M. Sánchez-Vizcaíno, L. Mur, and B. Martínez-López. "African swine fever: an epidemiological update". In: *Transboundary and emerging diseases* 59.s1 (2012), pp. 27–35.

[132] J.M. Sánchez-Vizcaíno, L. Mur, and B. Martínez-López. "African swine fever (ASF): five years around Europe". In: *Veterinary microbiology* 165.1 (2013), pp. 45–50.

[133] C.J. Sanders et al. "High-altitude flight of Culicoides biting midges." In: *Bulletin of Entomological Research* 94 (2011), pp. 123–136.

[134] T.C. Schroeder et al. "Economic impact of alternative FMD emergency vaccination strategies in the midwestern United States". In: *Journal of Agricultural and Applied Economics* 47.01 (2015), pp. 47–76.

[135] R.F. Sellers and P.S. Mellor. "Temperature and the persistence of viruses in Culicoides spp. during adverse conditions". In: *Revue scientifique et technique (International Office of Epizootics)* 12.3 (1993), 733—755.

[136] S. Siuli, Y. Li, and P. Wen. "Clustering technique-based least square support vector machine for EEG signal classification". In: *Computer methods and programs in biomedicine* 104.3 (2011), pp. 358–372.

[137] S. Solomos et al. “Density currents as a desert dust mobilization mechanism”. In: *Atmospheric Chemistry and Physics* 12.22 (2012).

[138] C. Stauffer, W. Grimson, and L. Eric. “Adaptive background mixture models for real-time tracking”. In: *Computer Vision and Pattern Recognition, 1999. IEEE Computer Society Conference on.* Vol. 2. IEEE. 1999.

[139] A. Stegeman et al. “Quantification of the transmission of classical swine fever virus between herds during the 1997-1998 epidemic in The Netherlands”. In: *Preventive veterinary medicine* 42.3 (1999), pp. 219–234.

[140] A. Stegeman et al. “The 1997-1998 epidemic of classical swine fever in the Netherlands”. In: *Veterinary Microbiology* 73.2 (2000), pp. 183–196.

[141] J.A. Stegeman et al. “Rate of inter-herd transmission of classical swine fever virus by different types of contact during the 1997-8 epidemic in The Netherlands”. In: *Epidemiology and infection* 128.02 (2002), pp. 285–291.

[142] A.F. Stein et al. “NOAA’s HYSPLIT atmospheric transport and dispersion modeling system”. In: *Bulletin of the American Meteorological Society* 96.12 (2015), pp. 2059–2077.

[143] M.A. Stevenson et al. “InterSpread Plus: a spatial and stochastic simulation model of disease in animal populations”. In: *Preventive veterinary medicine* 109.1 (2013), pp. 10–24.

[144] “The soil particle size dependent emission parameterization for an Asian dust (Yellow Sand) observed in Korea in April 2002”. In: *Atmospheric Environment* 37.33 (2003), pp. 4625–4636.

[145] D. Thompson et al. “Economic costs of the foot and mouth disease outbreak in the United Kingdom in 2001”. In: *Revue scientifique et technique-Office international des epizooties* 21.3 (2002), pp. 675–685.

[146] M. Thrusfield. *Veterinary epidemiology.* Elsevier, 2013.

[147] R. Tibshirani, G. Walther, and T. Hastie. “Estimating the number of clusters in a data set via the gap statistic”. In: *Journal of the Royal Statistical Society: Series B (Statistical Methodology)* 63.2 (2001), pp. 411–423.

[148] A. Torre et al. “Assessing the risk of African swine fever introduction into the European Union by wild boar”. In: *Transboundary and emerging diseases* 62.3 (2015), pp. 272–279.

[149] M. Tsidulko et al. “Numerical study of a very intensive eastern Mediterranean dust storm, 13-16 March 1998”. In: *Journal of Geophysical Research: Atmospheres* 107.D21 (2002).

[150] F.A. Verhoef, G.J. Venter, and C.W. Weldon. “Thermal limits of two biting midges, Culicoides imicola Kieffer and C. bolitinos Meiswinkel (Diptera: Ceratopogonidae)”. In: *Parasites & vectors* 7.1 (2014), p. 384.

[151] S. Viazzi et al. “Image feature extraction for classification of aggressive interactions among pigs”. In: *Computers and Electronics in Agriculture* 104 (2014), pp. 57–62.

[152] P. Viola and M. Jones. "Rapid object detection using a boosted cascade of simple features". In: *Computer Vision and Pattern Recognition, 2001. CVPR 2001. Proceedings of the 2001 IEEE Computer Society Conference on*. Vol. 1. IEEE. 2001, pp. I–511.

[153] Valderrama W. "Evaluación económica y epidemiológica en programas de salud animal: el caso de fiebre aftosa en Perú". In: (2010).

[154] K. Wadhwa and D. Wright. *A survey of Technology for the Elderly*. 2010.

[155] J.C. Warner et al. "Development of a three-dimensional, regional, coupled wave, current, and sediment-transport model". In: *Computers & Geosciences* 34.10 (2008), pp. 1284–1306.

[156] *WHO strategic action plan for pandemic influenza*. Tech. rep. World Health Organisation, 2007.

[157] A. Wilson and P. Mellor. "Bluetongue in Europe: vectors, epidemiology and climate change". In: *Parasitology Research* 103.1 (2008), pp. 69–77.

[158] A.J. Wilson and P.S. Mellor. "Bluetongue in Europe: past, present and future". In: *Philosophical Transactions of the Royal Society of London B: Biological Sciences* 364.1530 (2009), pp. 2669–2681.

[159] E.J. Wittmann, P.S. Mellor, and M. Baylis. "Effect of temperature on the transmission of orbiviruses by the biting midge, Culicoides sonorensis". In: *Medical and veterinary entomology* 16.2 (2002), pp. 147–156.

[160] S. Wortmann et al. "A new analytical approach to simulate the pollutant dispersion in the PBL". In: *Atmospheric Environment* 39.12 (2005), pp. 2171–2178.